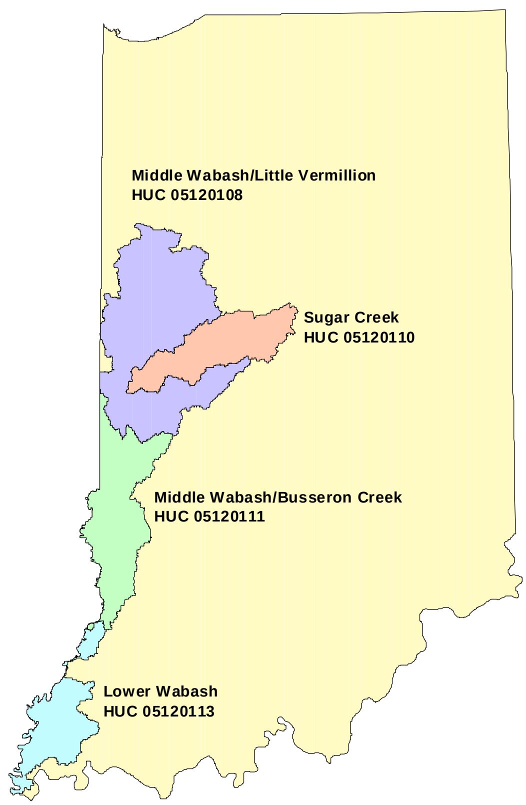
1999 Paired Site Study
Lower Wabash River Basin
Indiana



Indiana Department of Environmental Management
Office of Water Quality
Assessment Branch
Surveys Section

IDEM 032/02/035/2001
March 2001

1999 Paired Site Study

Lower Wabash River Basin

Indiana

By

Carl C. Christensen
Environmental Manager

Geographical Information Systems
Jim W. Butler

Compilation and development of the final report was the primary responsibility of the Surveys Section
Arthur C. Garceau, Surveys Section Chief

Indiana Department of Environmental Management
Office of Water Quality
Assessment Branch
Surveys Section
IDEM 032/02/035/2001
March 2001

When citing this document:

Christensen C. 1999. *1999 Paired Site Study, Lower Wabash River Basin, Indiana.* Indiana Department of Environmental Management, Office of Water Quality, Assessment Branch, Surveys Section, Indianapolis, Indiana.
IDEM 032/02/035/2001

ABSTRACT

The *1999 Paired Site Study, Lower Wabash River Basin, Indiana* was conducted to determine if water quality parameters changed significantly as the water flowed from a probabilistically selected site to the nearest downstream public access point, usually a bridge. The probabilistic sites were compared with the nearest downstream access points, which are referred to as “paired sites”.

Water chemistry samples and field measurements were taken at 34 paired sites in the Lower Wabash River Basin from June 6, 1999, to September 28, 1999. The data were analyzed using both parametric and non-parametric statistical tests, as well as a practical review of the data. Twenty-eight chemical and field parameters were examined using both statistical methods and practical review. The project hypothesized there was no significant change in water chemistry parameters comparing the probabilistic sites to the paired sites. The alternative hypothesis was there was a significant change in water chemistry comparing the probabilistic sites to the paired sites. The data were examined by analyzing the entire set of sampling sites and two subsets of sites; the mainstem Wabash River sites, and the lower order stream sites. Both parametric and nonparametric statistical test methods consistently had the same results.

Total solids were statistically different for the entire set of sites, the subset of lower order stream sites, and the subset of the mainstem Wabash River sites. When the Wabash River sites were tested as a subset, chloride and hardness were also determined to be statistically different comparing probabilistic to paired sites. Some parameters were not statistically tested because most of their observations were below the reporting limit. A practical review of these parameters did not disclose discernable patterns, although lower reporting limits could have changed this perception.

This type of program appears to be more applicable to smaller streams due to their relatively close proximity to downstream access points where the affects of point and non-point sources are minimized or avoided. The conclusions for this study indicate sampling at the nearest downstream public access sites during 1999 in the Lower Wabash River Basin was a viable alternative to sampling at probabilistically generated sites. Further studies are recommended to confirm the findings of this investigation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF PLATES

Page Intentionally Left Blank

INTRODUCTION

In 1996, the Surveys Section, Assessment Branch, Office of Water Quality, began using probabilistically selected sites to determine the overall surface water quality for all of the basins in the state of Indiana. These sites were provided by the United States Environmental Protection Agency's Environmental Management and Assessment Program (EMAP) through the staff of USEPA National Health and Environmental Effects Research Laboratory, Corvallis, Oregon. EMAP sites provide a statistically valid method of assessing the overall surface water quality within the area being studied.

Staff in the Surveys Section had historically sampled at public access locations, usually bridges. Sampling at bridges is believed to bias results because of their proximity to roads, anthropogenic activities, and the dumping of trash and debris into the streams.

A study was proposed in 1999 to determine if various water quality parameters changed as the water flowed from the probabilistically generated sites to the nearest point of public access downstream. In other words, the probabilistic site would form a "pair" with a site at the nearest downstream access, usually a bridge. Although two basins were studied by the Assessment Branch in 1999, the Kankakee River Basin and the Lower Wabash River Basin, this study used sites selected in only the Lower Wabash River Basin. This was done to reduce variability and provide greater control to the study. The sampling of the probabilistic sites with the paired sites were also conducted throughout the sampling season. This was done so observations would reflect various flow stages and weather conditions which occurred during the sampling season.

MATERIALS AND METHODS

Probabilistic sites were generated by USEPA in the Lower Wabash River Basin. These sites were pre-surveyed to determine if the site had sufficient water, gain access from landowners, verify safety, and map access routes. Once the probabilistic sites were determined, a simple random number generator was used to select subsets of these sites for pairing.

Probabilistic sites that were selected for pairing were projected on maps and the nearest downstream access was visually identified. This location became known as the "paired site". An initial goal of sampling 40 pairs of sites was set. In one case, two probabilistic sites on the Wabash River had the same downstream paired location. These were to be spread throughout the sampling season with 13 being sampled in May and June, 14 being sampled in July and August, and 13 being sampled in September and October. Table 1 lists the sites that were successfully paired and sampled.

Table 1 Probabilistic and Paired Sites Sampled for the Study

EMAP Code	Paired Site	Probabilistic Site	Date Sampled	Waterbody	Distance Apart (Miles)
134-001	WLV190-0006	WLV190-0004	9/28/99	Big Raccoon Creek	3.80
134-004	WBU150-0001	WBU150-0003	9/29/99	Turtle Creek	0.49
134-006	WSU060-0001	WSU060-0005	6/15/99	Sugar Mill Creek	0.08
134-017	WLV050-0001	WLV050-0002	6/9/99	Mud Pine Creek	1.04
134-025	WSU030-0001	WSU030-0003	6/8/99	Armentrout Dredge Ditch	0.18
134-026	WSU020-0001	WSU020-0005	10/18/99	Sugar Creek	1.16
134-029	WLV170-0001	WLV170-0004	8/10/99	Big Raccoon Creek	0.12
134-030	WLV040-0001	WLV040-0004	9/15/99	Vanatta Ditch	0.26
134-042	WBU030-0001	WBU030-0003	6/16/99	Otter Creek	0.56
134-045	WLV090-0001	WLV090-0004	6/9/99	Wabash River	4.40
134-046	WSU020-0002	WSU020-0006	6/8/99	Prairie Creek	2.16
134-052	WLW010-0001	WBU200-0001	9/15/99	Wabash River	8.17
134-053	WLV080-0001*	WLV080-0006	7/27/99	Wabash River	2.43
134-054	WSU010-0001	WSU010-0003	8/11/99	Browns Wonder Creek	1.55
134-057	WLV180-0001	WLV180-0010	8/10/99	Little Raccoon Creek	0.94
134-060	WLW100-0001	WLW080-0002	8/3/99	Wabash River	3.66
134-068	WBU040-0001	WBU040-0004	9/15/99	Wabash River	1.77
134-069	WLV080-0002*	WLV070-0002	6/9/99	Wabash River	15.55
134-086	WLV010-0001	WLV010-0004	9/15/99	Burnett Creek	1.19
134-089	WLV060-0001	WLV060-0002	6/9/99	Fall Creek	0.42
134-092	WBU160-0001	WBU160-0004	6/21/99	Busseron Creek	1.16
134-093	WLV180-0002	WLV180-0011	9/28/99	Williams Creek	0.15
134-096	WLW080-0001	WLW060-0001	6/17/99	Wabash River	19.84
134-098	WSU010-0002	WSU010-0004	6/8/99	Sugar Creek	0.62
134-101	WSU060-0002	WSU060-0008	8/4/99	Sugar Creek	0.17
134-106	WBU030-0002	WBU030-0004	7/28/99	Otter Creek	1.06
134-114	WLV180-0003	WLV180-0012	9/22/99	Little Raccoon Creek	0.18
134-121	WSU050-0001	WSU050-0007	7/27/99	Black Creek	0.20
134-130	WLV040-0002	WLV040-0005	6/8/99	Big Pine Creek	0.32
134-141	WSU060-0003	WSU060-0009	7/28/99	Roaring Creek	1.17
134-142	WSU030-0002	WSU030-0005	7/26/99	Little Potato Creek	1.17
134-145	WLV030-0001	WLV030-0007	9/21/99	Wabash River	9.43
134-149	WLV090-0002	WLV090-0005	6/9/99	Spring Creek	0.33
134-161	WLV190-0002	WLV190-0005	8/4/99	Big Raccoon Creek	1.20

*These two sites are the same geographic location.

Sampling methods for chemical analytes and field measurements followed procedures outlined in *Field Procedures Manual*, 1998 (Beckman and Hall, 1998). Samples at paired sites were not taken directly from the bridge, but were taken by wading into the stream upstream of the bridge. For two sets of sites; WLW010-0001 with WBU200-0001, and WBU040-0001 with WBU040-0004, the samples were taken by boat. Samples were collected at the paired site before the probabilistic site. This was done to ensure that sediments that would be disturbed at the probabilistic site would not be mixed into the water column and interfere with samples taken at the paired site downstream. The following tables list the field parameters and the chemical analytes that were measured.

Table 2 Field Parameters

Parameter	Method	Reporting Limit
Dissolved Oxygen	SM 4500-OG	0.03 mg/L
Turbidity	SM 2130	0.3 NTU
Specific Conductance	SM 2510	3 umhos/cm
Temperature	SM 2550	-5° Celsius
PH	SM 4500-H	+/-0.01 SU

Table 3 Chemical Parameters for Laboratory Analyses

Anions/Physical			Nutrients/Organic		
Parameter	**MTD**	**CRQL**	**Parameter**	**MTD**	**CRQL**
Alkalinity	310.1	10 mg/L	Total Kjeldahl Nitrogen (TKN)	351.2	.05 mg/L
Total Solids	160.3	1.0 mg/L	Ammonia-N	350.1	.01 mg/L
Suspended Solids	160.2	4.0 mg/L	Nitrate+Nitrite-N	353.2	.01 mg/L
Dissolved Solids	160.1	1.0 mg/L	Total Phosphorus	356.2	1.0 mg/L
Sulfate	375.2	1.0 mg/L	Total Organic Carbon (TOC)	415.1	1.0 mg/L
Chloride	325.2	1.0 mg/L	Carbonaceous Oxygen Demand (COD)	410.4	3.0 mg/L
Hardness	130.1	1.0 mg/L	Cyanide – Total	335.3	.005 mg/L

Metals		
Parameter	**MTD**	**CRQL**
Arsenic	206.2	4.0 ug/L
Cadmium	213.2	1.0 ug/L
Chromium	218.2	3.0 ug/L
Copper	200.7	3.0 ug/L
Lead	239.2	2.0 ug/L
Mercury	245.1	0.2 ug/L
Nickel	249.2	2.0 ug/L
Selenium	270.2	1.0 ug/L
Zinc	200.7	10.0 ug/L

Statistical tests were conducted to determine if the water quality parameters were significantly different from the probabilistic sites to the paired sites. The null hypothesis was that there was no statistical change in water quality parameters when comparing the probabilistic sites to the paired sites. The alternative hypothesis was that there was a statistical change in water quality comparing the probabilistic sites to the paired sites.

Two statistical tests were performed: the Paired t-Test for dependent samples and the Wilcoxon Matched Pairs test. The Paired t-Test is strongest for normal shaped distributions while the Wilcoxon Matched Pairs test is a nonparametric test that can be applied to both normal and non-normal distributions. Data transformation was also explored for the Paired t-Test.

RESULTS AND DISCUSSION

Sampling Success and Limitations

Some of the 1999 probabilistic sites, chosen with a random number generator for this study, could not be paired. Twelve of the randomly selected probabilistic sites were ineligible for this study because the waterbody ended before it flowed under a downstream bridge. Three probabilistic sites selected by the random draw were essentially at bridge locations making them inappropriate for the study. One paired site was dropped during sampling because the location was unsafe.

The weather conditions also made sampling more difficult. Drought conditions during 1999 resulted in some of the waterbodies becoming dry. Although a goal was set of sampling 40 paired sites, only 34 paired sites were successfully sampled. Thirteen sets of sites were sampled from May to June, 11 sets of sites were sampled from July to August, and 10 sets of sites were sampled from September to October.

Spatial Relationship of Sites

The mean distance between all of the pairs of sites in the study was 2.56 miles with a standard deviation of 4.42 miles. This large standard deviation indicates a great deal of variability. The instance where the probabilistic and paired sites were furthest apart was on the Wabash River for sites WLW080-0001 and WLW060-0001 which were 19.84 miles apart. The closest a pair of sites were to each other were sites WSU060-0001 and WSU060-0005 on Sugar Mill Creek. These sites were only 0.08 miles apart

The large variability for the entire set of sites was a result of the large distances between the probabilistic Wabash River sites and their respective paired sites. Due to the relative scarcity of bridges, the sites on the Wabash River were much further apart compared to sites on the lower order streams. The mean distance between the probabilistic sites and their respective pairs on the Wabash River was 8.16 miles with a standard deviation of 6.55 miles. This can be compared to the mean distances of the lower order streams, which had a mean distance of 0.83 miles with a standard deviation of only 0.81 miles.

Statistical Tests

The data for the parameters were first examined to determine if statistical testing was appropriate for the data set. Some of the parameters had "sparse data" where there were few or no actual values above the detection limit. In these cases, statistical tests can be conducted, but are not meaningful. These parameters will be discussed in a following section.

For the paired t-test, the null hypothesis was the mean numeric difference between the paired site parameters and the probabilistic site parameters was 0 and the alternate hypothesis was the mean numeric difference between the paired site parameters and the probabilistic site parameters was not equal to 0. This is mathematically expressed as:

$$H_o\text{:}\ \ \Delta X=0$$
and
$$H_a\text{:}\ \ \Delta X<>0.$$

The test was 2-tailed and the confidence level for the test was 95%. The results of these tests are in Table 4. The null hypothesis was rejected when the p value was less than 0.05.

Table 4 Paired t-Test Results

Parameter	p Values For All Sites n=34	p Values for Lower Order Stream Sites n=26	p Values For Wabash River Sites n=8
Alkalinity	0.846	0.647	0.929
Chloride	0.179	0.264	0.017*
COD	0.251	0.498	0.104
Conductivity	0.104	0.308	0.140
Dissolved Oxygen**	0.761	0.666	0.543
Hardness	0.376	0.068	0.011*
Nitrate + Nitrite	0.700	0.211	0.390
TKN	0.351	0.164	0.582
PH	0.836	0.660	0.478
Total Phosphorus	0.708	0.809	0.715
Total Dissolved Solids	0.258	0.165	1.000
Total Suspended Solids	0.844	0.225	0.353
Total Solids	0.011*	0.024*	0.010*
Sulfate	0.689	0.656	0.257
Temperature	0.822	0.481	0.186
TOC	0.781	0.924	0.501
Turbidity	0.362	0.520	0.284
Zinc	0.889	0.782	0.671

* Null hypothesis rejected.

** One dissolved oxygen result on the Wabash River was rejected due to equipment failure.

For all observations, the null hypothesis was rejected for only total solids. However, the two components of total solids, total dissolved solids and total suspended solids, were not statistically different. Figures 1 through 3 graphically show the observations of these three parameters.

Figure 1 Total Solids Concentrations for Probabilistic and Paired Sites

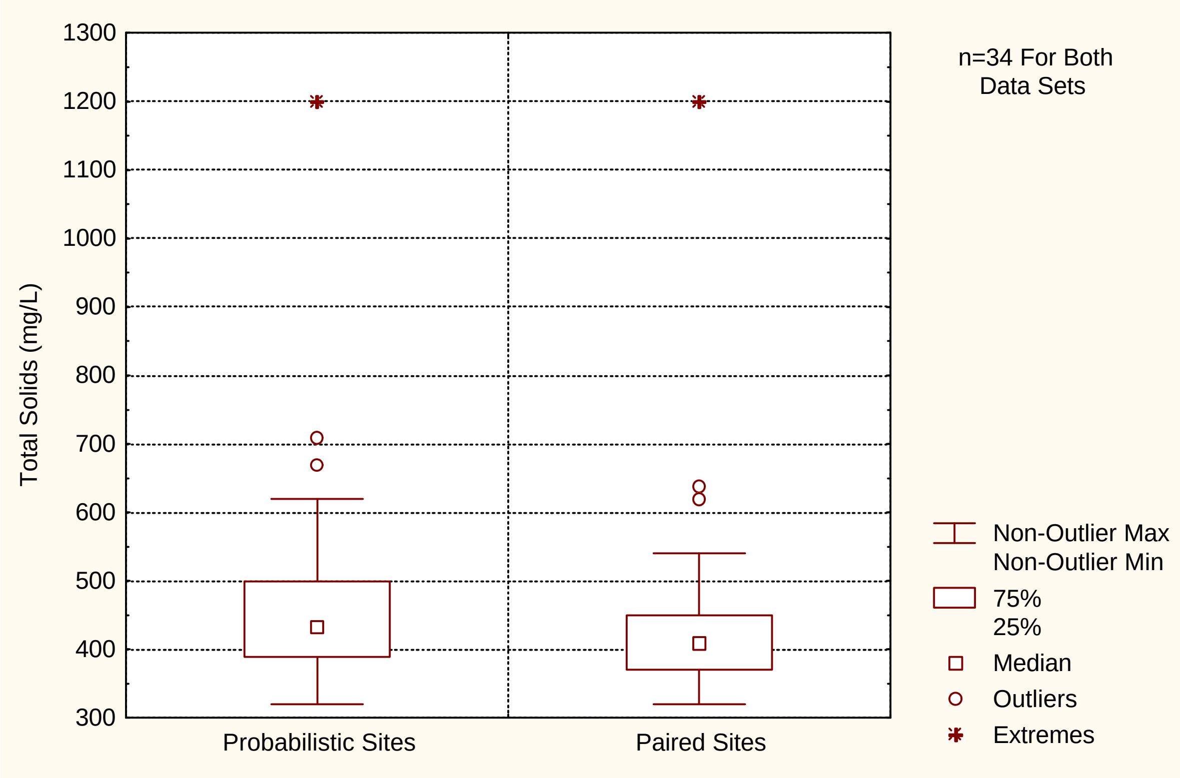


Figure 2 Total Suspended Solids for Probabilistic and Paired Sites

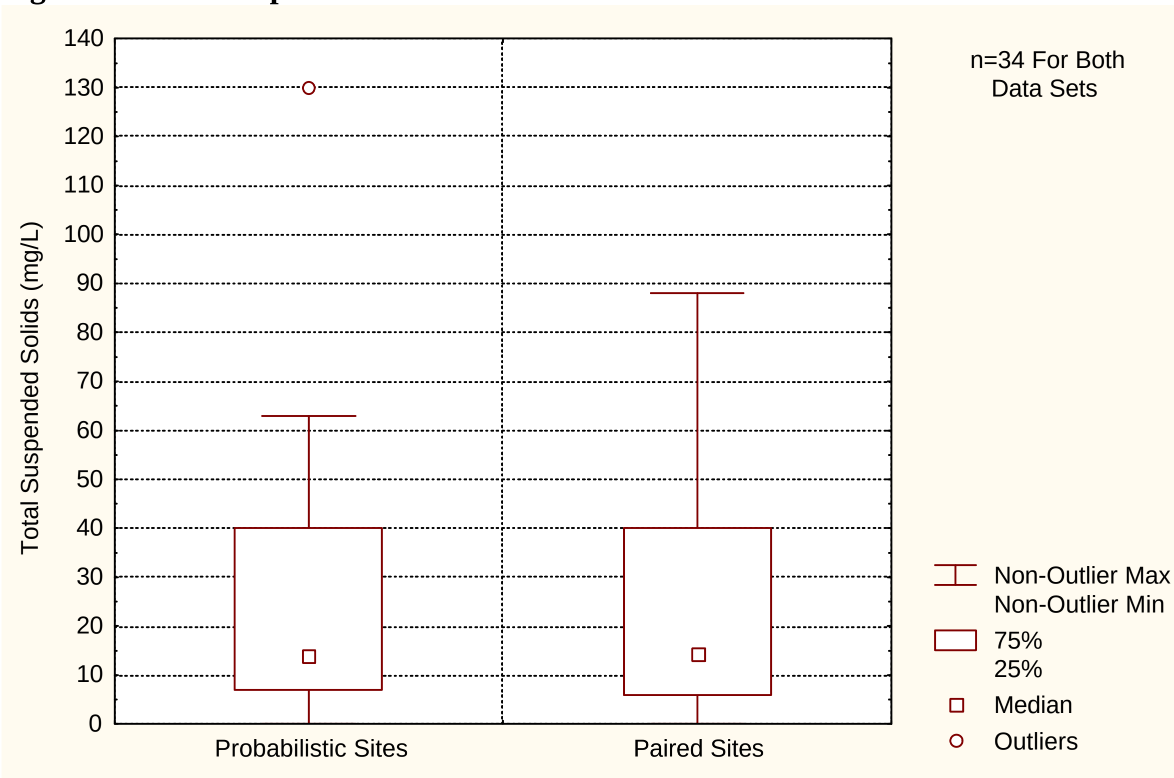
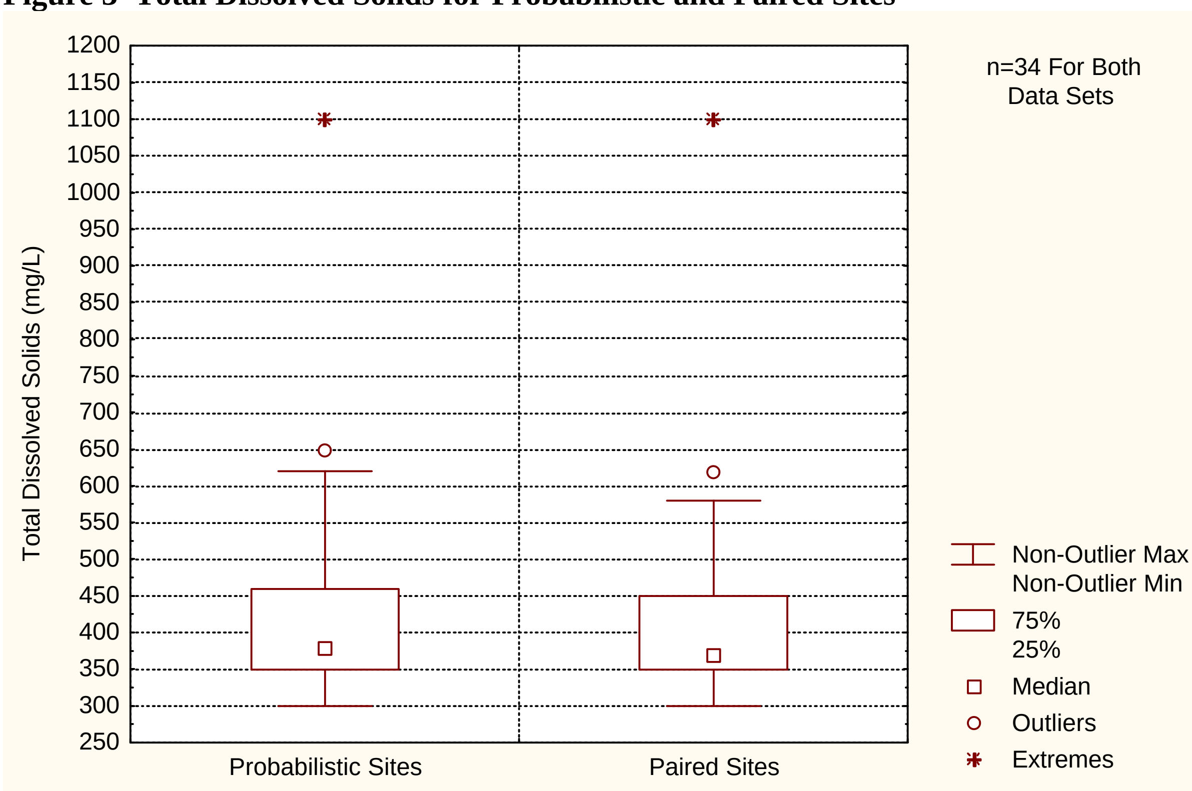


Figure 3 Total Dissolved Solids for Probabilistic and Paired Sites



The standard deviation for total solids is a combination of the standard deviations of both the dissolved and suspended solids. This cumulative effect is the probable cause of why the null hypothesis was rejected for total solids.

Sites near and at bridges tend to be deeper due to dredging and channelization, creating slower water. When the water in streams slows down, particulate matter settles onto the sediments. Therefore, paired sites would be expected to have lower amounts of organic and inorganic particulates in the water. However, nutrients associated with particulate matter, TOC, COD, TKN, and total phosphorus were not statistically different. Suspended solids and turbidity, which includes both inorganic and organic material, were also not statistically different.

In general, it is expected that there would be less canopy cover at bridge sites compared to probabilistic sites. Canopy cover directly affects the amount of photosynthesis which results in increasing dissolved oxygen and pH. Dissolved oxygen and pH were not statistically different at the paired sites compared to the probabilistic sites.

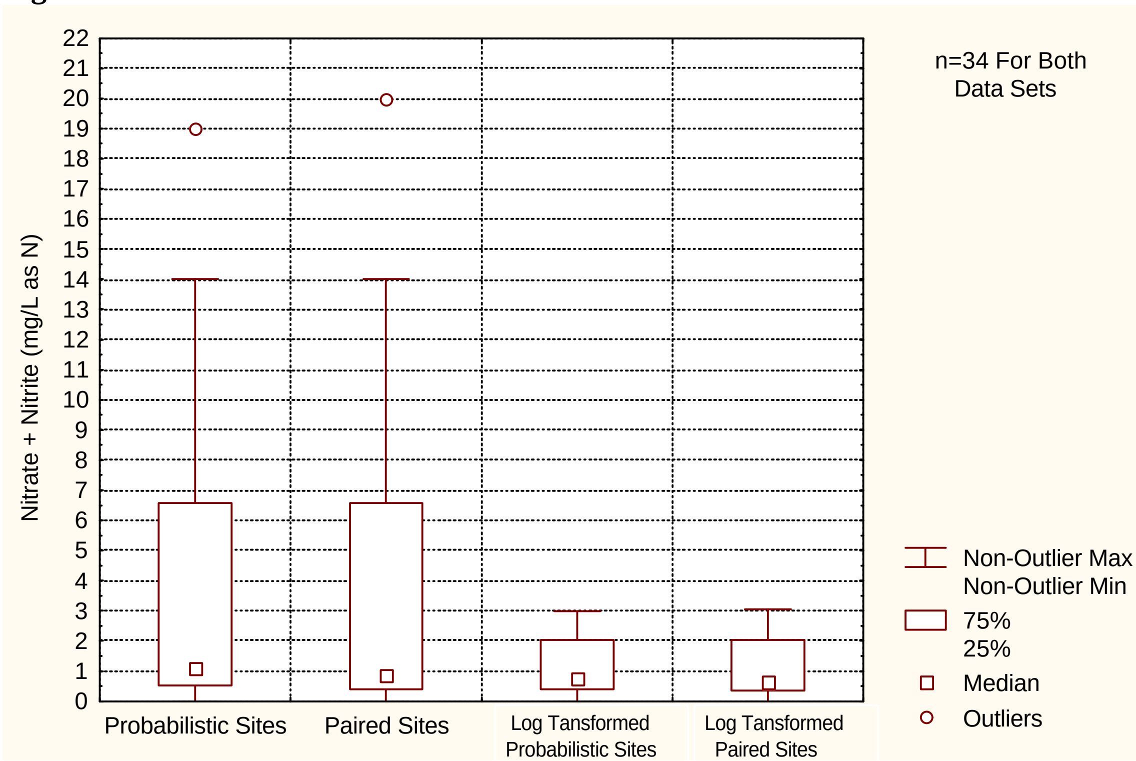
For the subset of mainstem Wabash River sites, two parameters rejected the null hypothesis in addition to total solids. These were hardness and chloride. This is probably the result of the relatively large distances between the paired and probabilistic sites which allow more inputs from both point and nonpoint sources.

The Paired t-Test is most robust when the data has a normal shape. Outliers and skewness of the data set reduce the statistical power of parametric tests like the Paired t-Test. The data for the tests in Table 4 were further explored by logarithm transformation.

The data was transformed by adding 1.0 to the observation and taking the natural logrithm of the sum. Adding 1.0 to the observation is required because observations less than the reporting limit were initially assigned a value of 0.0, and the natural logarithm cannot be taken of 0.0. In all but one case, the decision to reject or not reject the null hypothesis was the same as with the non-transformed data.

The exception was the nitrate + nitrite Paired t-Test for the lower order streams. Logarithm transformed data resulted in a p value of only 0.023, which is low enough to reject the null hypothesis. However, the test on the transformed data may not be any more robust than the non-transformed data. Figure 4 shows the box-whisker plots for the nitrate + nitrite observations in question.

Figure 4 Nitrate + Nitrite Observations for Lower Order Streams



Although the outliers are removed by log transformation, the median is still in the lower portion of the box, indicating skewness. The log transformed data is represented by histograms in Figures 5 and 6.

Figure 5 Log Transformed Nitrate + Nitrite Results for Lower Order Prob. Sites

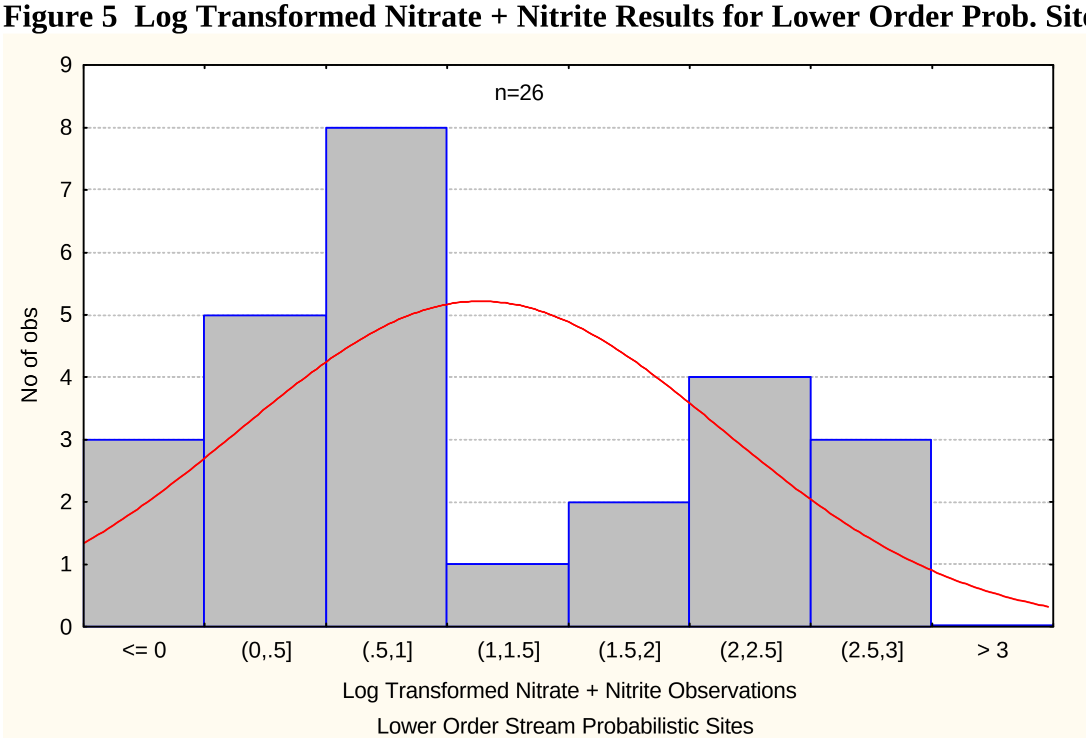
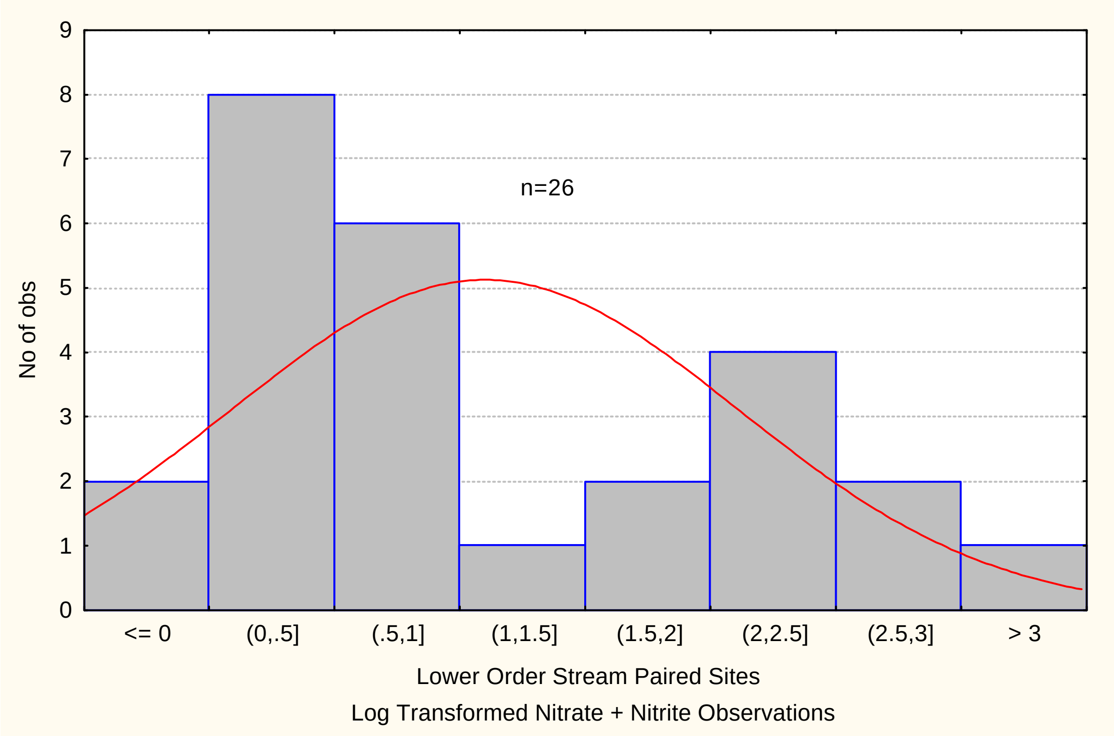


Figure 6 Log Transformed Nitrate + Nitrite Results for Lower Order Paired Sites



In both cases, the data sets do not have a normal appearance. Both sets appear to have a bi-modal appearance.

Two additional nitrate + nitrite transformations were performed for the lower order probabilistic and paired sites. These were taking the natural logrithm of the observations + 0.1 and the natural logrithm of the observations + 0.01. In these two cases, the null hypothesis was not rejected for the lower order streams, with p values of 0.05 and 0.48, respectively. The appearance of these data sets still remained skewed.

Non-parametric testing has the advantage of not being dependent on the shape of the data sets. The results of this type of test are a good way of resolving the contradiction created by the log transformation process, and further verifying the results of the Paired t-Test, which is generally regarded as a powerful test.

The data were tested using the non-parametric Wilcoxon Matched Pairs Test, with a null hypothesis that the two sets of populations, the probabilistic sites and the paired sites, were the same. The results are listed below in Table 5. The null hypothesis is rejected when the p value is less than 0.05. The expressions describing the null and alternative hypothesis are:

H_o: Paired and Probabilistic observations came from the same population.

H_a: Paired and Probabilistic observations came from different populations.

Table 5 Wilcoxon Matched Pairs Test Results

Parameter	p Values For All Sites n=34	p Values For Lower Order Stream Sites n=26	p Values For Wabash Sites n=8
Alkalinity	0.611	0.936	0.353
Chloride	0.165	0.937	0.018*
COD	0.304	0.794	0.107
Conductivity	0.469	0.809	0.141
Dissolved Oxygen**	0.860	0.882	0.753
Hardness	0.891	0.141	0.018*
Nitrate + Nitrite	0.184	0.064	0.686
TKN	0.526	0.435	0.834
PH	0.844	0.753	0.779
Total Phosphorus	0.248	0.201	0.779
Total Dissolved Solids	0.242	0.235	0.675
Total Suspended Solids	0.641	0.289	0.499
Total Solids	0.001*	0.030*	0.017*
Sulfate	0.681	0.509	0.123
Temperature	0.713	0.316	0.208
TOC	0.955	0.955	0.889
Turbidity	0.503	0.205	0.575
Zinc	0.902	0.808	1.000

* Null Hypothesis Rejected
** One dissolved oxygen result on the Wabash River was rejected due to equipment failure.

Although the Wilcoxon Matched Pairs test is regarded to be less powerful than the Paired t-Test, the conclusions from the tests are the same. The only parameter that was found to be statistically different for all of the sites was total solids. The conclusions for the two subsets, the Wabash River sites and the lower order streams, were the same as the Paired t-Test. This agreement validates the results of the both statistical test methods, although nitrate + nitrite test for lower order streams had a p value of only 0.064. Lower reporting limits or more samples may have resulted in different conclusions for this test.

Comparison of Parameters not Statistically Tested

Arsenic, ammonia, cadmium, chromium, cyanide, lead, mercury, nickel, and selenium had infrequent observations above the reporting limit. Although the selected statistical tests could be performed, the results would not be meaningful. However, this does not preclude a practical look at the data. Table 6 lists the number of observations above the detection limit for the probabilistic sites, the paired sites, and the number of occasions where both sites were above the reporting limit.

Table 6 Observations Above the Detection Limit

Parameter	Probabilistic Sites	Paired Sites	Simultaneous Detections
Ammonia	13	13	10
Arsenic	6	3	2
Cadmium	0	0	0
Chromium	1	1	1
Copper	17	11	7
Cyanide	0	1	0
Lead	7	10	5
Mercury	0	0	0
Nickel	7	7	5
Selenium	7	5	5

Cadmium and mercury had no observations above the detection limit. From a practical standpoint, there was no discernable difference between the paired site and the probabilistic site for these two parameters, although lower reporting limits could have resulted in different conclusions. Cyanide had one observation above the reporting limit at site WSU050-0001 on Black Creek. This observation was very low, 0.014 mg/L. Since this value was near the reporting limit, it may be considered similar to the result at the probabilistic site.

Chromium had only one paired and probabilistic site with observations above the reporting limit. This was at sites WLW080-0001 and WLW060-0001 on the Wabash River with concentrations of 5.8 and 6.9 ug/L, respectively. Considering the proximity of these measurements to the reporting limit, 3.0 ug/L, and the low concentrations observed, these measurements can be considered very similar.

Arsenic, lead, nickel, and selenium had reporting limits of 4.0, 2.0, 2.0, and 1.0 ug/L, respectively. There were a few cases for each of these parameters when one of the two sites had an observation above the reporting limit and the other did not. The largest difference between a pair of sites where one had an observation above the reporting limit and the other site in the set did not were sites WLV180-0001 and WLV180-0010 on Little Raccoon Creek. In this case, the paired site, WLV180-0001 had a lead observation of 4.1 ug/L while the probabilistic site, WLV180-0010, was below the reporting limit of 2.0 ug/L.

Ammonia and copper had the highest number of observations above the reporting limit that are listed in Table 6. Although the Paired t-test is not appropriate, the Wilcoxon Matched Paired test can still be done, with the caveat that lower reporting limits could have resulted in different conclusions. These results are in Table 7.

Table 7 Wilcoxon Matched Pairs Test Results for Sparse Data Parameters

Parameter	p Value for All Sites n=34	p Values Excluding Wabash Sites n=26	p Values For Wabash Sites n=8
Ammonia	0.650	0.959	0.423
Copper	0.330	0.394	0.866

The null hypothesis was not rejected for any of the above cases. Therefore, the null hypothesis that ammonia and copper concentrations are the same comparing probabilistic and paired sites is accepted.

INTERPRETATION OF SUMMARY DATA AND GRAPHICS

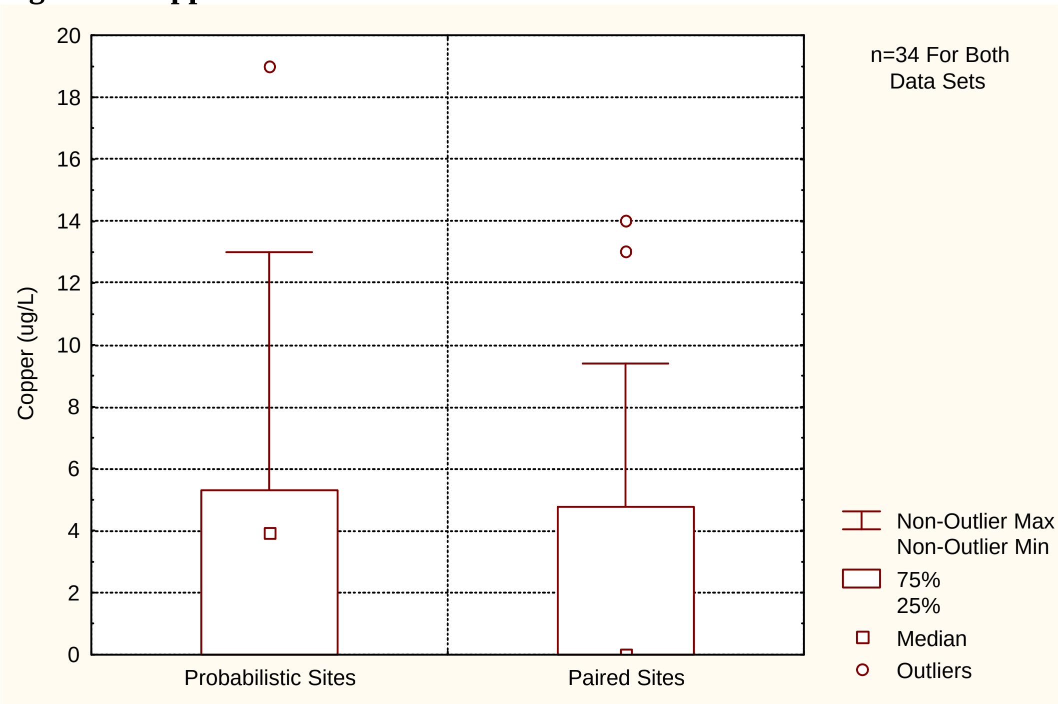
The data were plotted into box-whisker plots so the various parameters could be visually compared to each other. The box-whisker plots can be found in Appendix 1. Summary statistics for statistically tested parameters, including copper and ammonia, can be found in Appendix 2.

The data sets are not independent of each other. The water at the probabilistic sites influences the concentrations observed at the downstream paired sites. The expectation is there would be very little visual difference between the data sets for the probabilistic and paired sites.

Overall, there was very little visual difference comparing the box-whisker plots of the probabilistic and paired sites for most of the parameters. Most of the plots in Appendix 1 show virtually no difference, or very slight increases or decreases between median values and quartile ranges. Examination of the summary statistics in Appendix 2 show very similar median, mean, and 95% confidence interval ranges when comparing the probabilistic to paired observation sets. However, there are exceptions to this case.

In Figure 7, the box-whisker plots for copper show very different median results comparing the two sets of data. The 95% confidence interval was also noticeably different with a range of 5.14 to 2.05 for the probabilistic sites and 3.96 to 1.08 for the paired sites. This was a result of the large number of observations below the detection limit for copper at the paired sites. As mentioned in the previous section, these sets were not statistically different comparing the Paired to Probabilistic sites when using the Wilcoxon Matched Pairs Test.

Figure 7 Copper Concentrations for Probabilistic and Paired Sites



The mean turbidity results for the probabilistic sites was 55.5 NTU compared to a much lower mean of 28.8 NTU for the paired sites. Although this suggests a large decrease in turbidity as the water travels from upstream to downsteam, the outliers and extreme outliers can be censored and the trimmed means can be reexamined. When the means are calculated without outliers, the results are more similar with mean values at 19.3 and 25.1 for the probabilistic and paired sites, respectively.

CONCLUSIONS AND RECOMMENDATIONS

The intent of the probabilistic program is to determine the overall water quality of the basins in the state of Indiana and locate streams and rivers that are impaired for water quality. Access to probabilistic sites is not always certain. Factors that prohibit access to probabilistic sites include: landowner denial, physical barriers, and unsafe conditions.

The results of this study strongly suggest that sampling at downstream access sites when possible will not affect the overall picture of water quality in Indiana. When sampling at the probabilistic site is not possible, sampling at the nearest downstream bridge may be an option that can be utilized. This is most applicable to lower order streams where the distance to the nearest downstream bridge is relatively low.

Caution must be utilized when and if this is done. This study was performed only in one basin, the Lower Wabash River Basin. The results for a similar study in a different basin could be different or could further validate the findings in this study. Different results may be realized if the stream levels and weather conditions are not the same as experienced in 1999.

The cornerstone of scientific investigation are reproducible scientific experiments. Duplication of the results of this study in other basins with more sampling sites would further validate the results, or show basins and conditions where the significant statistical difference can be determined at paired sites. Further studies should be done to confirm the findings in this report.

SUMMARY

The 1999 Paired Study was conducted to determine if significant changes occurred between probabilistic sites and the nearest downstream public access location. This was usually a bridge and was referred to as a “paired site”. Thirty-four probabilistic sites were successfully matched with paired sites in this study.

The data were analyzed using parametric and non-parametric tests. The data were divided into three groups: the entire set of sites, the Wabash River sites, and the lower order stream sites. For all three data sets, total solids was statistically different comparing the probabilistic sites to the paired sites. When the data were also tested for the subset of Wabash River sites, hardness and chloride were also statistically different. The greater variation in the Wabash River subset was attributed to the relatively large distance in miles between the probabilistic and paired sites compared to the lower order streams. The larger distances allow for contributions to the water from various point and nonpoint sources. The resulting analysis for nitrate + nitrite concentrations are somewhat inconclusive for the subset of lower order streams, although based on the information in this study, no statistical difference exists between the probabilistic and paired sites.

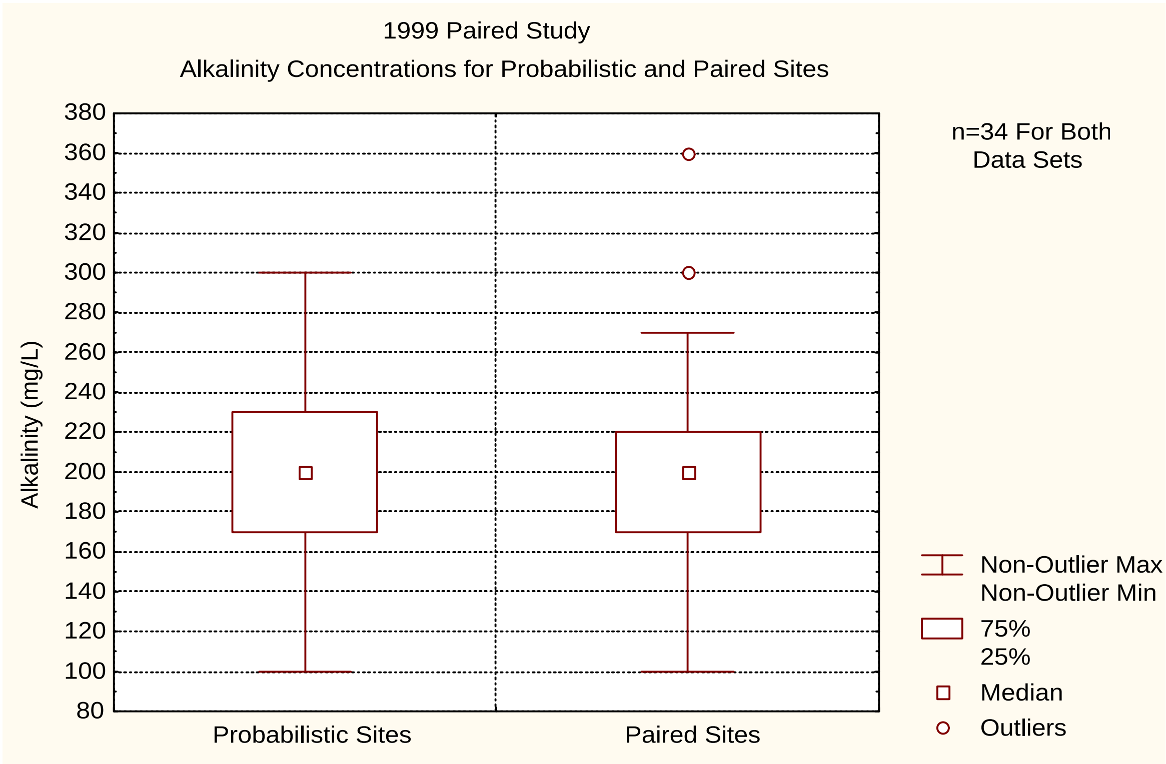
The data were also examined visually and parameters that could not be statistically tested were also reviewed. Although some box and whisker plots suggest very slight changes in the water chemistry between the probabilistic and paired sites, the results of the statistical tests suggest that an extremely large data set would be required to reject the null hypothesis, if indeed this could be accomplished. Interpretation of data with few observations above the reporting limit did not indicate discernable patterns.

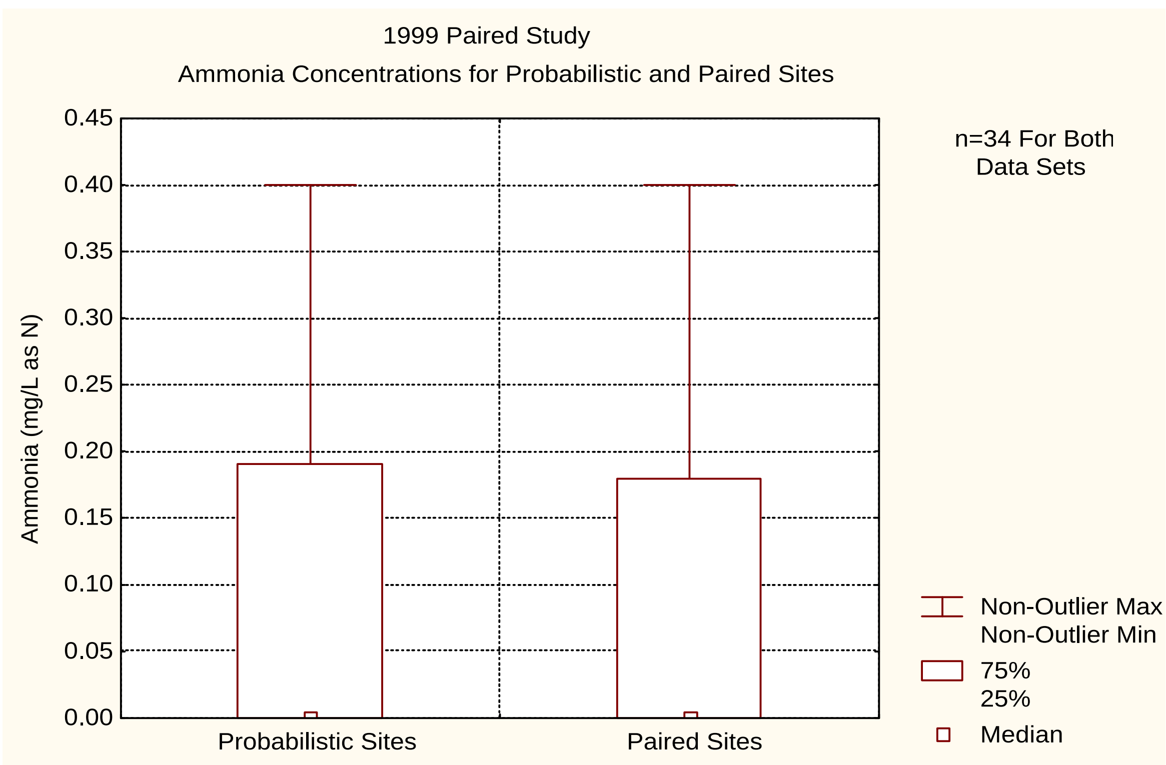
Access to probabilistic sites is not certain due to a variety of problems and obstacles. Although further studies should be done to confirm these findings and conclusions, this type of sampling may serve as an alternative to sampling directly at probabilistic sites.

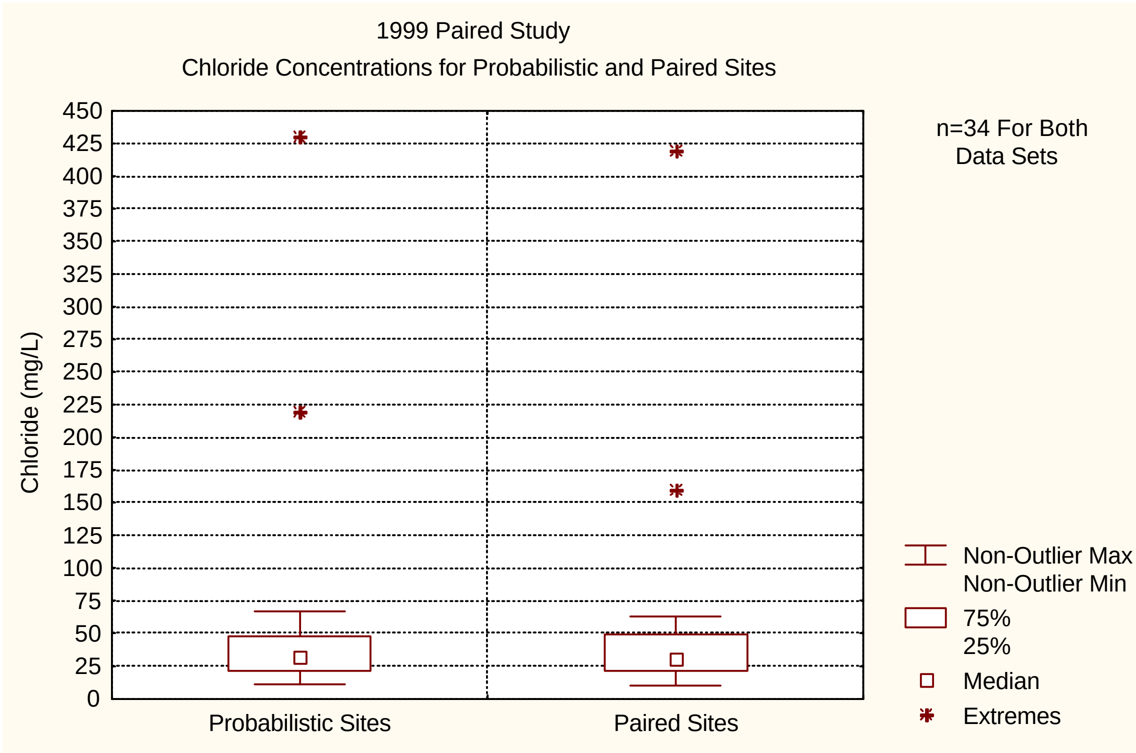
LITERATURE CITED

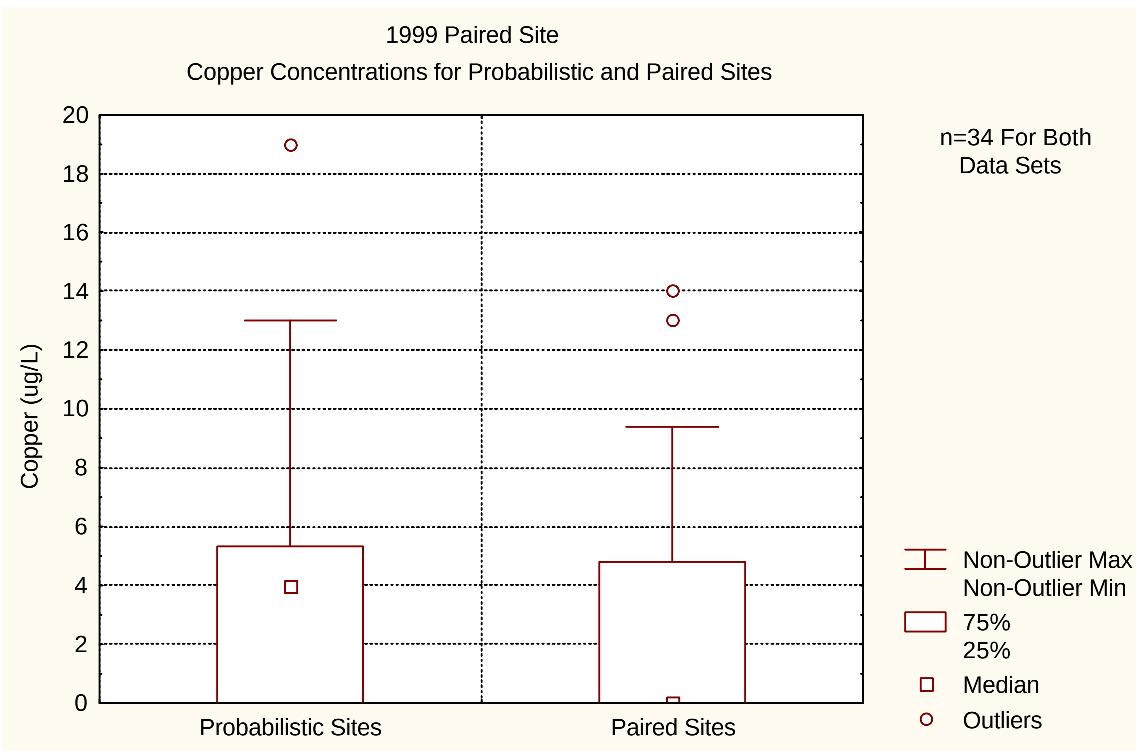
Beckman T, Hall S. 1998. *Field Procedure Manual, 1998*. Indiana Department of Environmental Management, Office of Water Management, Assessment Branch, Surveys Section, Indianapolis, Indiana. IDEM 032/02/012/1998

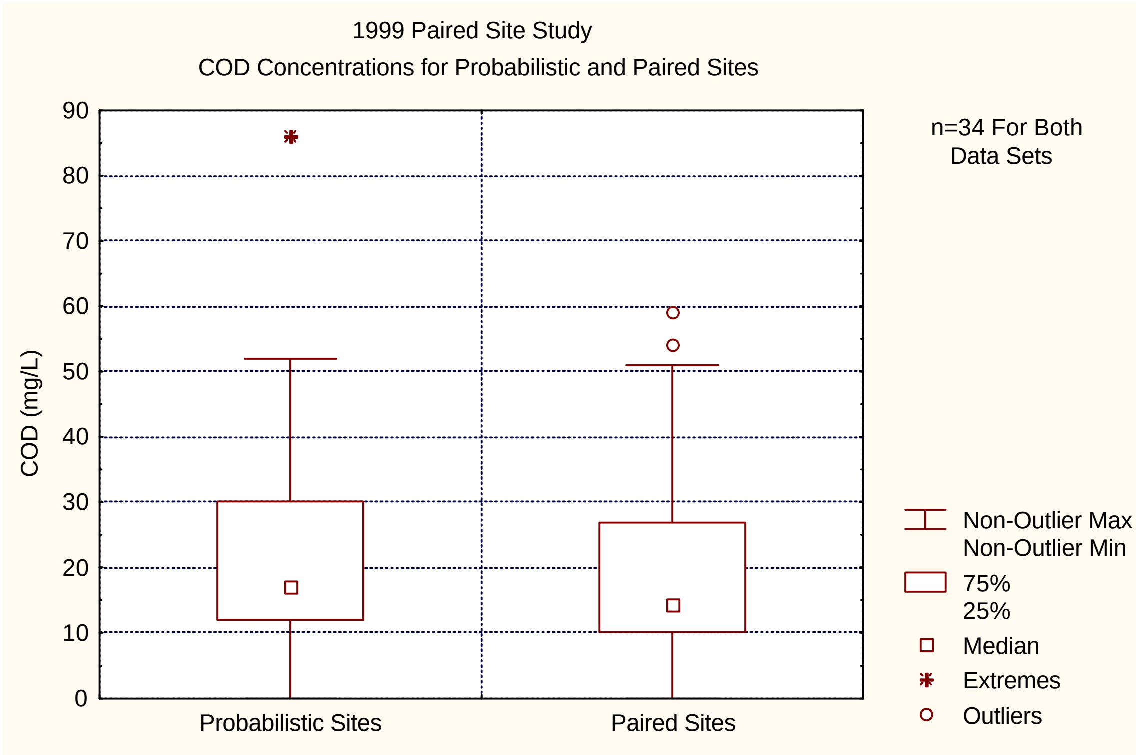
APPENDIX 1
Graphical Comparisons of Probabilistic and Paired Sites

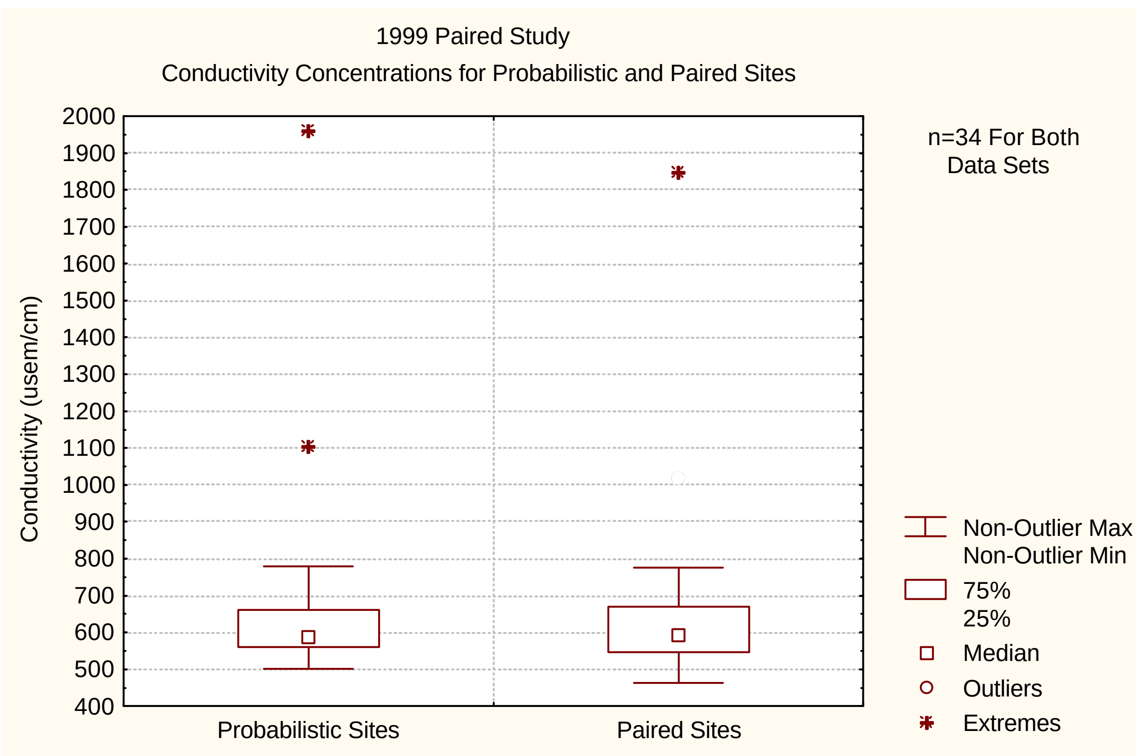


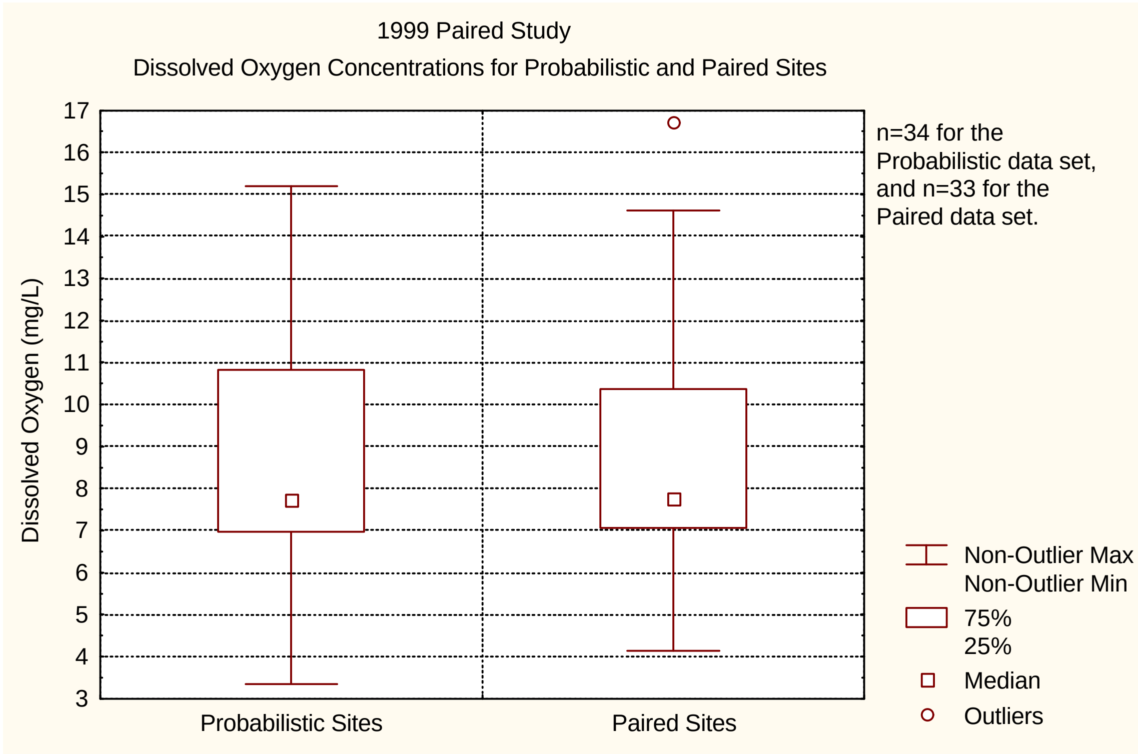


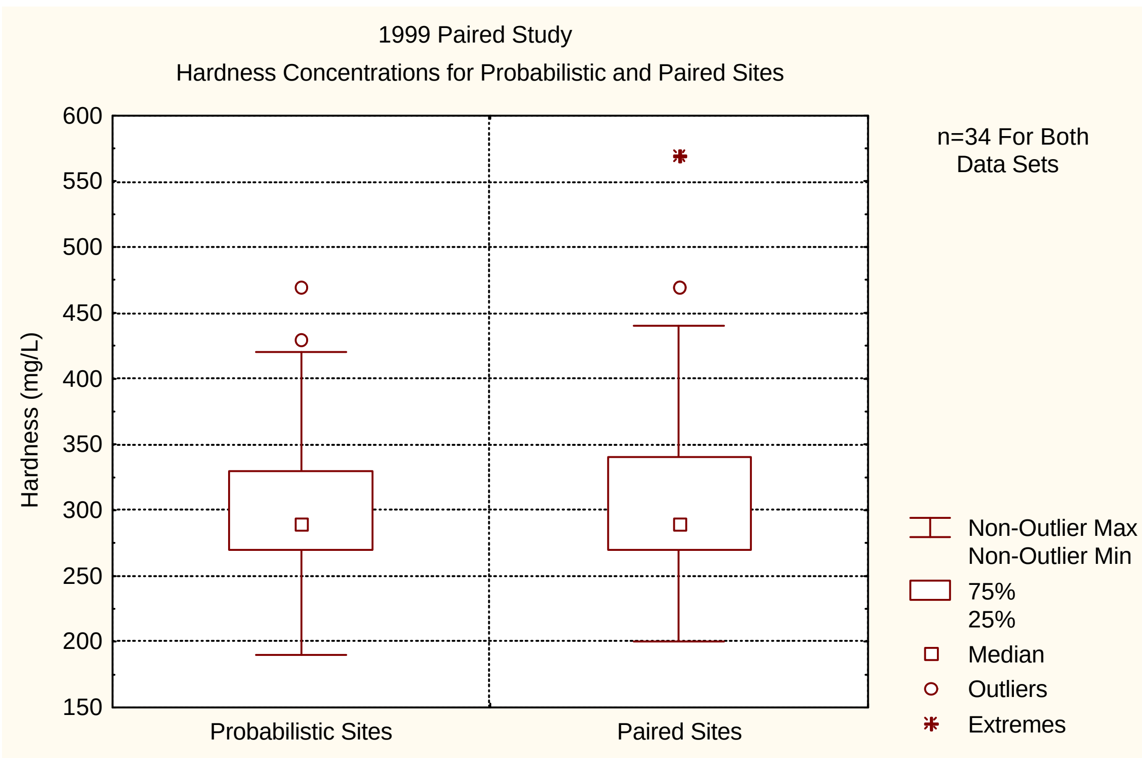

1999 Paired Study
Chloride Concentrations for Probabilistic and Paired Sites
Chloride (mg/L)
450
425
400
375
350
325
300
275
250
225
200
175
150
125
100
75
50
25
0
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Extremes

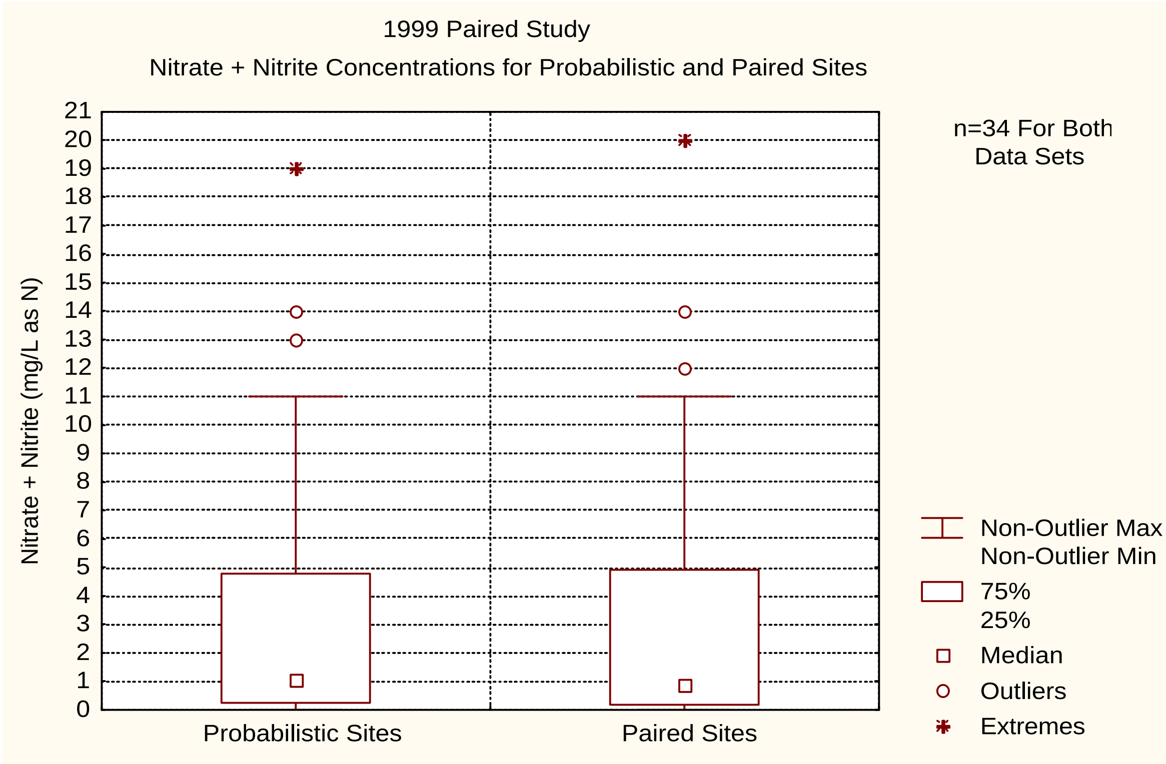

1999 Paired Site
Copper Concentrations for Probabilistic and Paired Sites
Copper (ug/L)
20
18
16
14
12
10
8
6
4
2
0
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers

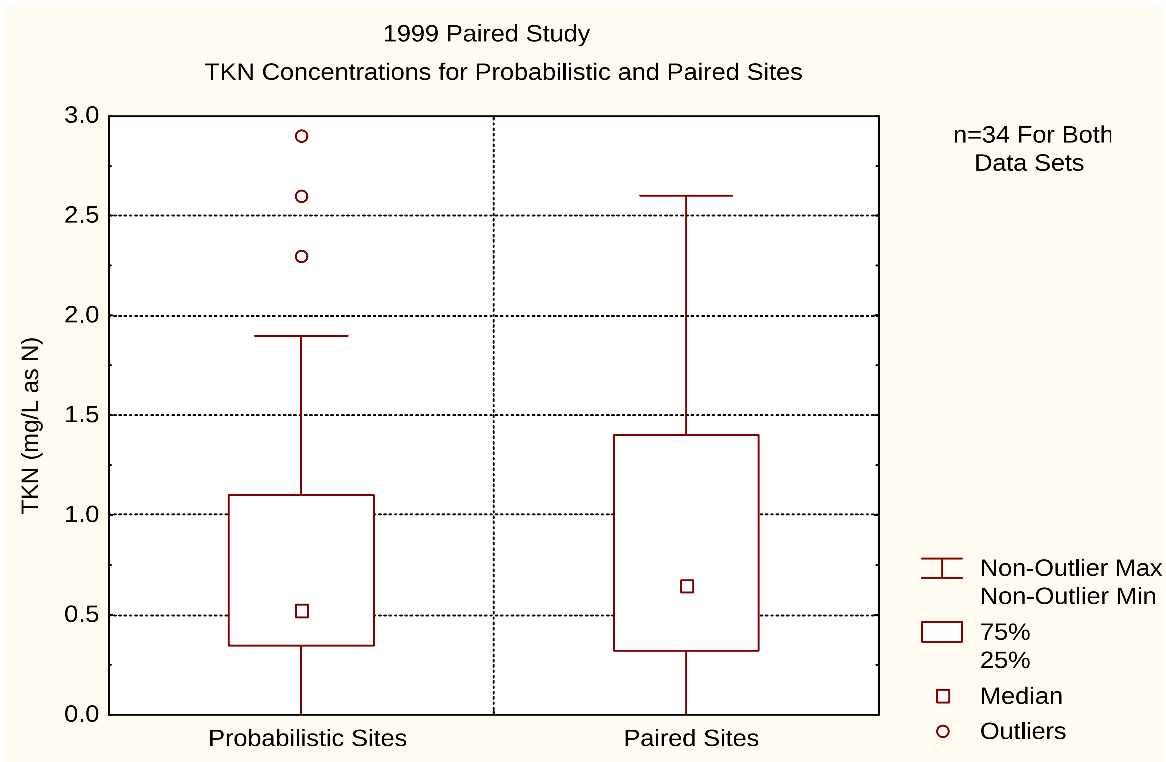

1999 Paired Site Study
COD Concentrations for Probabilistic and Paired Sites
COD (mg/L)
0
10
20
30
40
50
60
70
80
90
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Extremes
Outliers

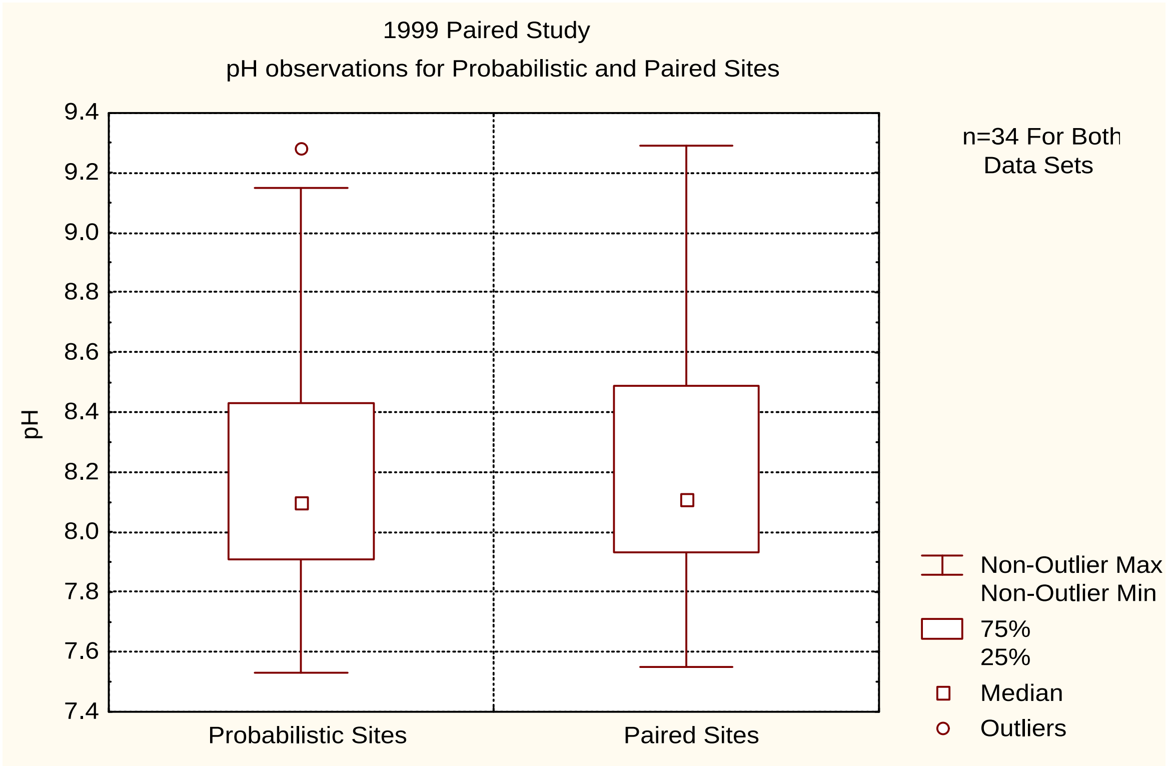

1999 Paired Study
Conductivity Concentrations for Probabilistic and Paired Sites
Conductivity (usem/cm)
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers
Extremes

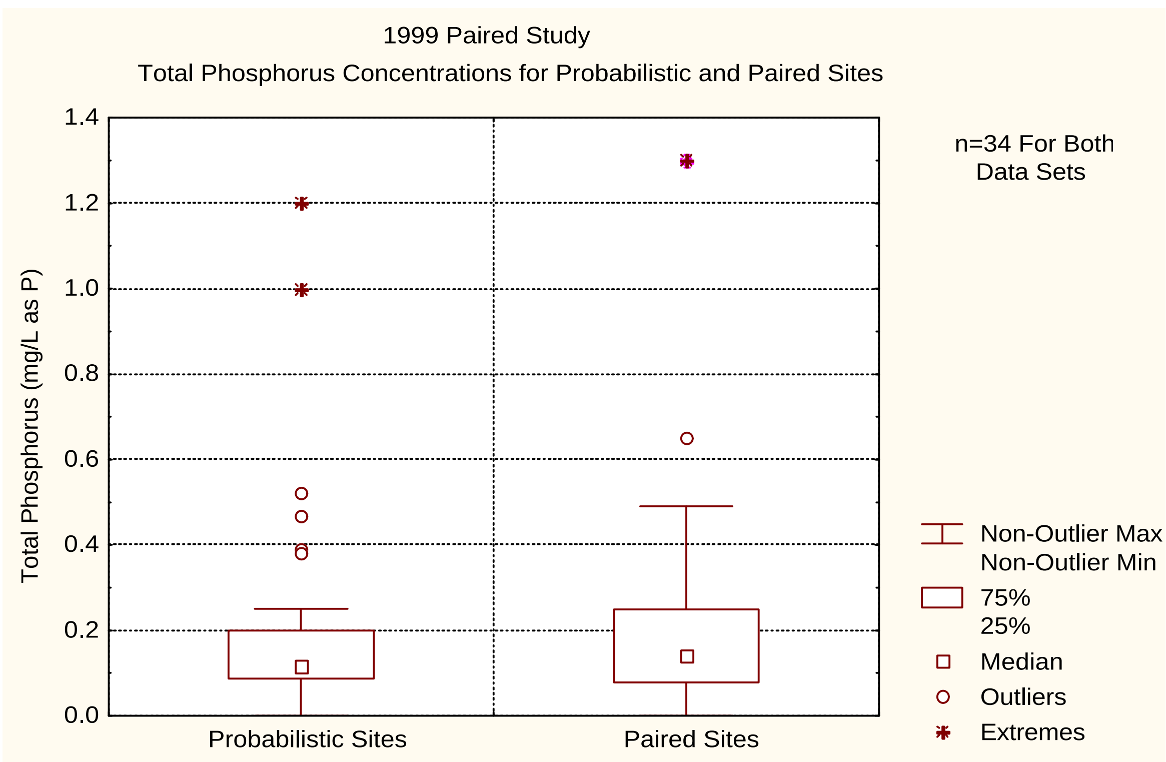

1999 Paired Study
Dissolved Oxygen Concentrations for Probabilistic and Paired Sites
Dissolved Oxygen (mg/L)
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
Probabilistic Sites
Paired Sites
n=34 for the Probabilistic data set, and n=33 for the Paired data set.
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers

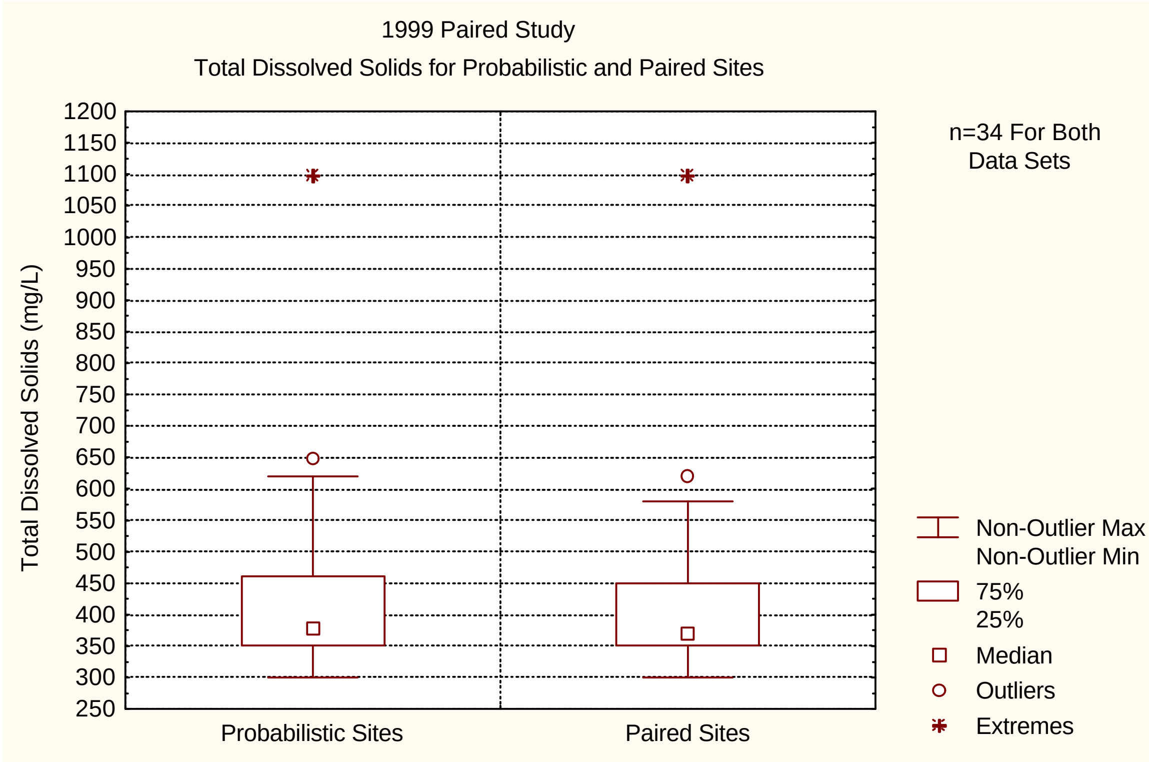

1999 Paired Study
Hardness Concentrations for Probabilistic and Paired Sites
Hardness (mg/L)
600
550
500
450
400
350
300
250
200
150
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers
Extremes

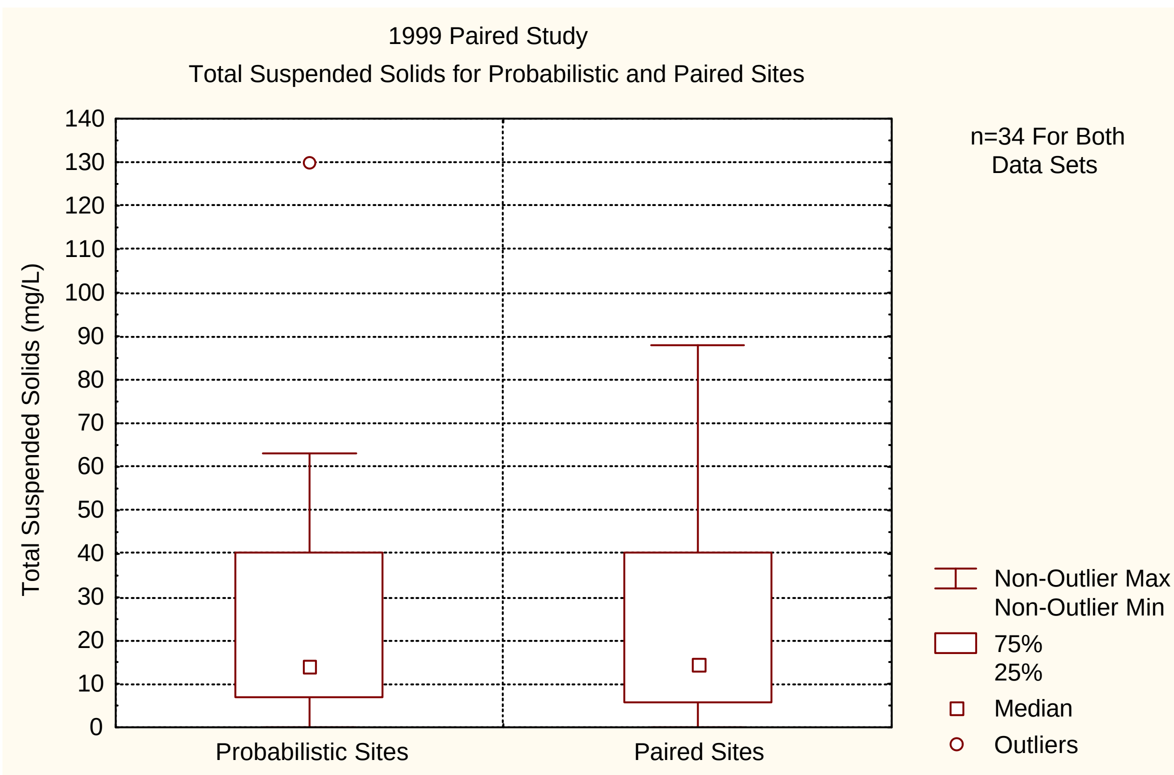

1999 Paired Study
Nitrate + Nitrite Concentrations for Probabilistic and Paired Sites
Nitrate + Nitrite (mg/L as N)
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers
Extremes

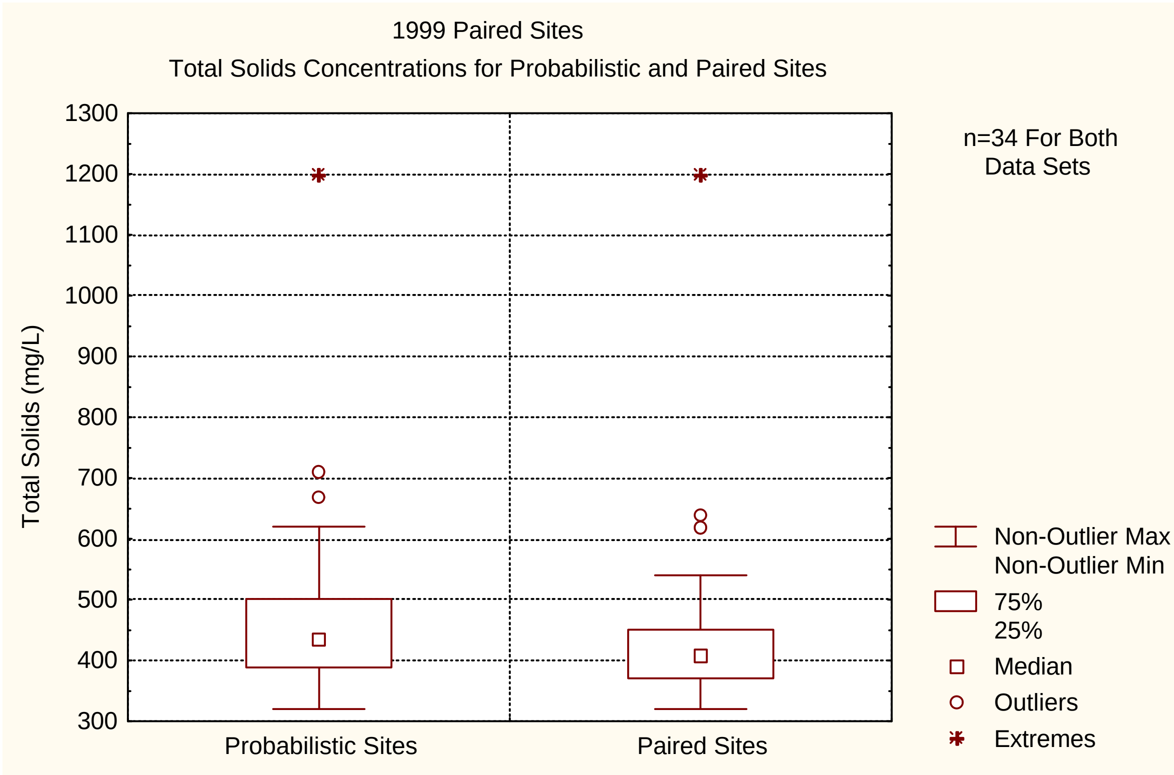

1999 Paired Study
TKN Concentrations for Probabilistic and Paired Sites
TKN (mg/L as N)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers

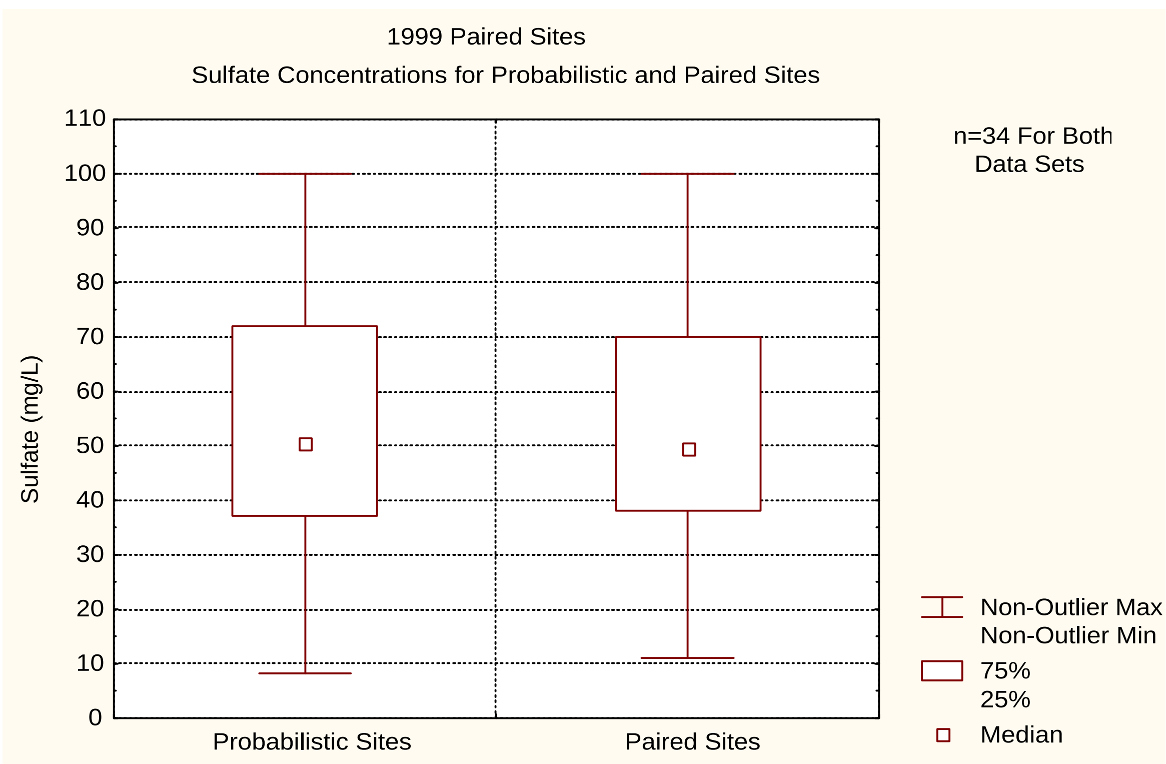

1999 Paired Study
pH observations for Probabilistic and Paired Sites
n=34 For Both Data Sets
pH
9.4
9.2
9.0
8.8
8.6
8.4
8.2
8.0
7.8
7.6
7.4
Probabilistic Sites
Paired Sites
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers

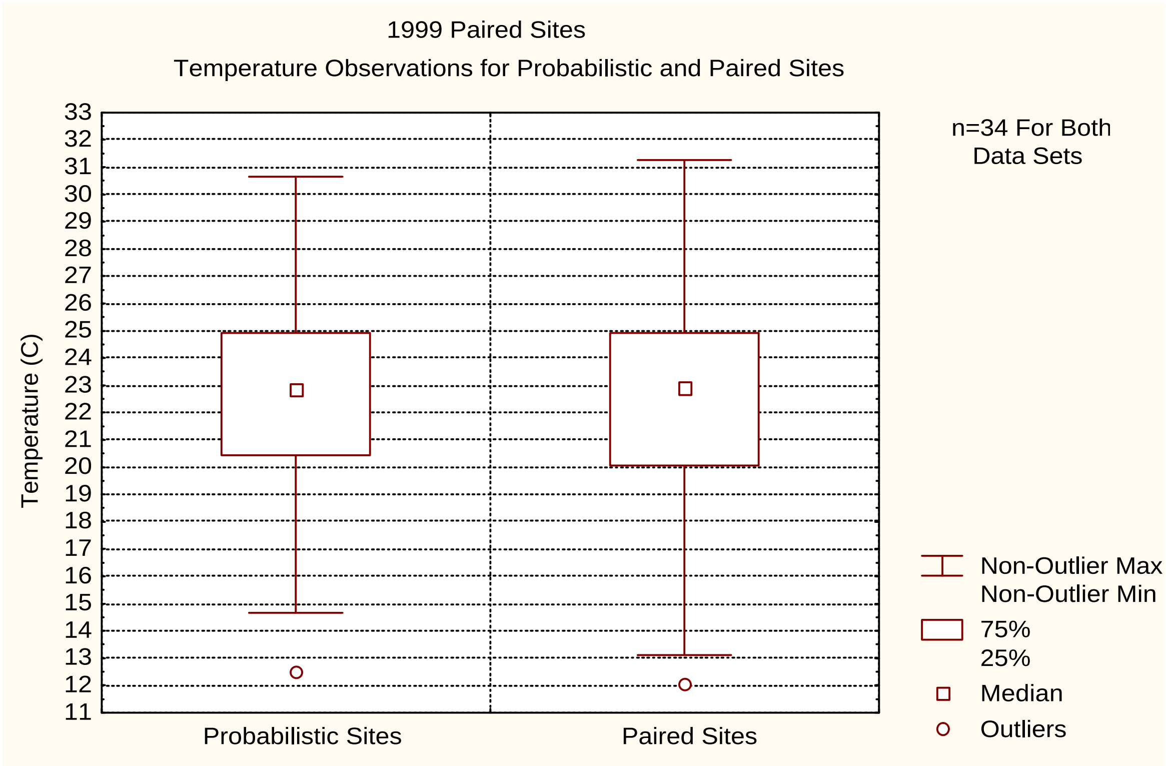

1999 Paired Study
Total Phosphorus Concentrations for Probabilistic and Paired Sites
n=34 For Both Data Sets
Total Phosphorus (mg/L as P)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Probabilistic Sites
Paired Sites
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers
Extremes

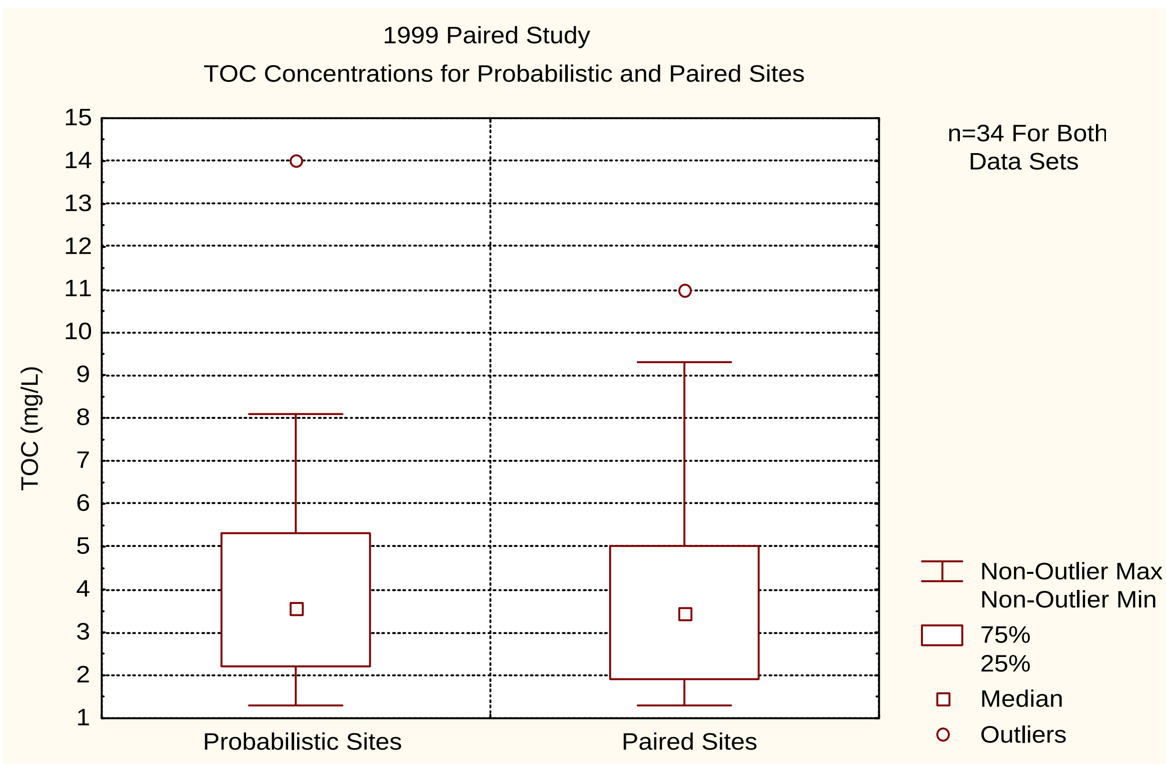

1999 Paired Study
Total Dissolved Solids for Probabilistic and Paired Sites
Total Dissolved Solids (mg/L)
1200
1150
1100
1050
1000
950
900
850
800
750
700
650
600
550
500
450
400
350
300
250
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers
Extremes

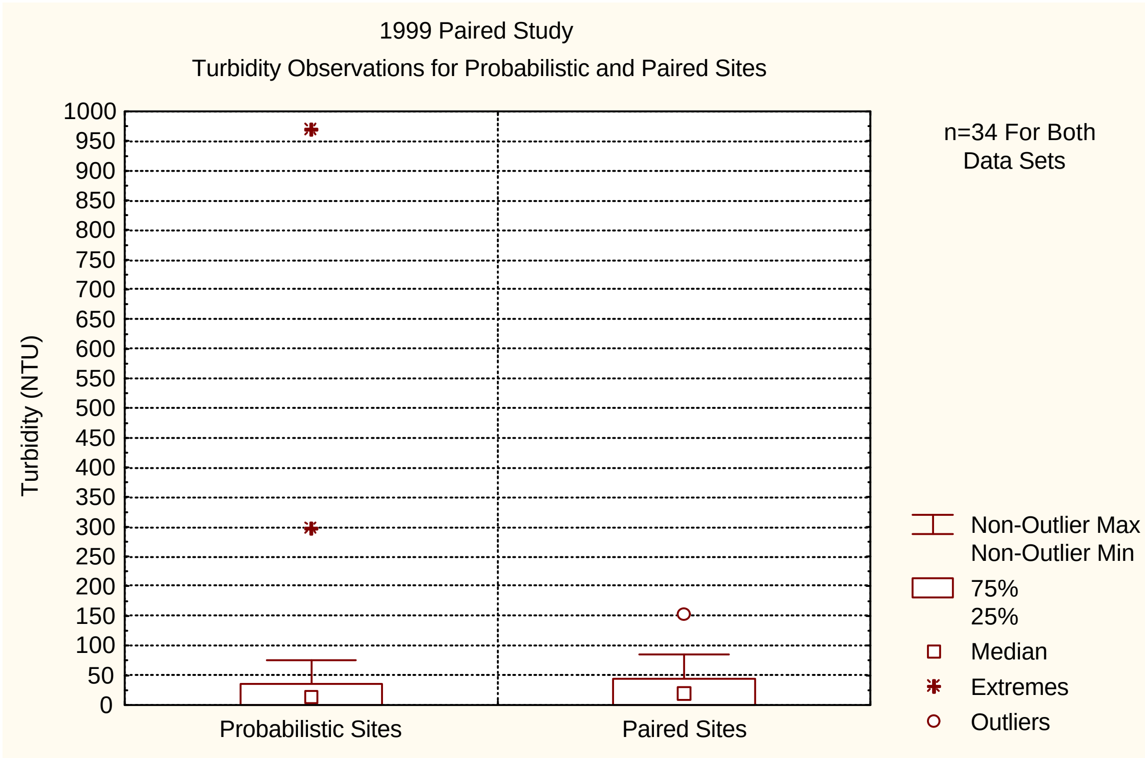

1999 Paired Study
Total Suspended Solids for Probabilistic and Paired Sites
Total Suspended Solids (mg/L)
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers

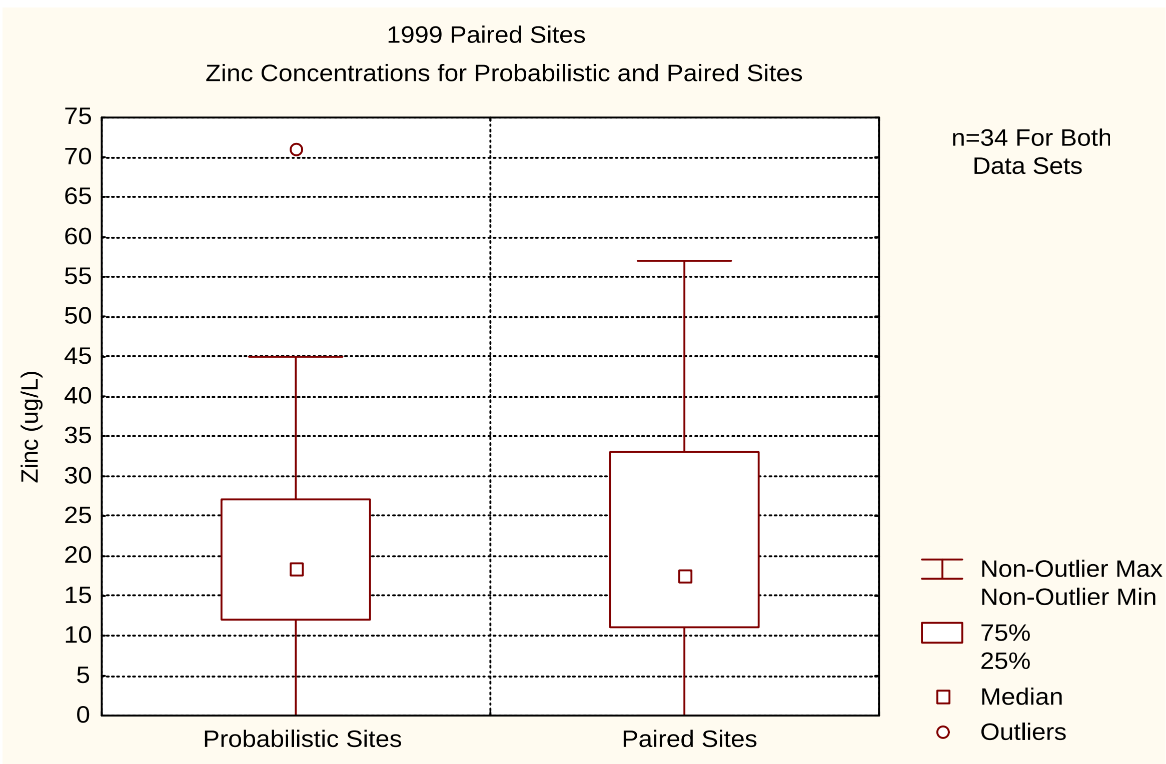

1999 Paired Sites
Total Solids Concentrations for Probabilistic and Paired Sites
Total Solids (mg/L)
1300
1200
1100
1000
900
800
700
600
500
400
300
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers
Extremes


1999 Paired Sites
Sulfate Concentrations for Probabilistic and Paired Sites
Sulfate (mg/L)
110
100
90
80
70
60
50
40
30
20
10
0
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median


1999 Paired Sites
Temperature Observations for Probabilistic and Paired Sites
Temperature (C)
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers


1999 Paired Study
TOC Concentrations for Probabilistic and Paired Sites
TOC (mg/L)
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers


1999 Paired Study
Turbidity Observations for Probabilistic and Paired Sites
Turbidity (NTU)
1000
950
900
850
800
750
700
650
600
550
500
450
400
350
300
250
200
150
100
50
0
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Extremes
Outliers


1999 Paired Sites
Zinc Concentrations for Probabilistic and Paired Sites
Zinc (ug/L)
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
Probabilistic Sites
Paired Sites
n=34 For Both Data Sets
Non-Outlier Max
Non-Outlier Min
75%
25%
Median
Outliers

APPENDIX 2
Summary Statistics of Probabilistic and Paired Sites

Parameter	n	Mean	- 95% Confidence Interval	+ 95% Confidence Interval	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Variance	Standard Deviation
Probabilistic Alkalinity (mg/L)	34	201.76	186.24	217.29	200.00	100.00	300.00	170.00	230.00	1978.61	44.48
Paired Alkalinity (mg/L)	34	200.29	182.89	217.70	200.00	100.00	360.00	170.00	220.00	2487.79	49.88
Probabilistic Ammonia (mg/L as N)	34	0.08	0.04	0.11	0.00	0.00	0.40	0.00	0.19	0.01	0.11
Paired Ammonia (mg/L as N)	34	0.07	0.04	0.11	0.00	0.00	0.40	0.00	0.18	0.01	0.10
Probabilistic Chloride (mg/L)	34	50.09	23.57	76.61	32.50	11.00	430.00	22.00	48.00	5775.78	76.00
Paired Chloride (mg/L)	34	47.62	22.90	72.33	31.00	10.00	420.00	22.00	49.00	5018.12	70.84
Probabilistic Copper (ug/L)	34	3.59	2.05	5.14	3.95	0.00	19.00	0.00	5.30	19.57	4.42
Paired Copper (ug/L)	34	2.52	1.08	3.96	0.00	0.00	14.00	0.00	4.80	17.07	4.13
Probabilistic COD (mg/L)	34	21.05	14.53	27.58	17.00	0.00	86.00	12.00	30.00	349.67	18.70
Paired COD (mg/L)	34	19.69	13.91	25.46	14.50	0.00	59.00	10.00	27.00	273.94	16.55
Probabilistic Conductivity (usem/cm)	34	662.32	573.35	751.30	591.00	501.00	1960.00	559.00	660.00	65023.07	255.00
Paired Conductivity (usem/cm)	34	651.53	569.16	733.90	596.50	463.00	1850.00	547.00	672.00	55733.77	236.08
Probabilistic D.O. (mg/L)	34	8.67	7.68	9.66	7.74	3.35	15.20	6.99	10.82	8.00	2.83
Paired D.O. (mg/L)	33	8.72	7.69	9.76	7.75	4.14	16.72	7.07	10.38	8.49	2.91
Probabilistic Hardness (mg/L)	34	305.88	283.17	328.59	290.00	190.00	470.00	270.00	330.00	4237.08	65.09
Paired Hardness (mg/L)	34	316.47	285.75	347.19	290.00	200.00	570.00	270.00	340.00	7750.80	88.04
Probabilistic Nitrate + Nitrite (mg/L as N)	34	3.56	1.83	5.30	1.08	0.00	19.00	0.27	4.80	24.75	4.98
Paired Nitrate + Nitrite (mg/L as N)	34	3.54	1.78	5.30	0.89	0.00	20.00	0.19	4.90	25.48	5.05
Probabilistic TKN (mg/L as N)	34	0.80	0.55	1.06	0.53	0.00	2.90	0.35	1.10	0.54	0.74
Paired TKN (mg/L)	34	0.87	0.62	1.12	0.65	0.00	2.60	0.32	1.40	0.51	0.71
Probabilistic pH	34	8.19	8.04	8.34	8.10	7.53	9.28	7.91	8.43	0.18	0.43
Paired pH	34	8.20	8.07	8.33	8.11	7.55	9.29	7.93	8.49	0.14	0.38
Probabilistic Total Phosphorus (mg/L as P)	34	0.21	0.12	0.30	0.12	0.00	1.20	0.09	0.20	0.07	0.26
Paired Total Phosphorus (mg/L as P)	34	0.20	0.12	0.29	0.14	0.00	1.30	0.08	0.25	0.06	0.24
Probabilistic Total Dissolved Solids (mg/L)	34	426.47	375.31	477.63	380.00	300.00	1100.00	350.00	460.00	21496.26	146.62
Paired Total Dissolved Solids (mg/L)	34	419.71	369.09	470.32	370.00	300.00	1100.00	350.00	450.00	21045.37	145.07
Probabilistic Total Suspended Solids (mg/L)	34	24.32	14.95	33.70	14.00	0.00	130.00	7.00	40.00	722.16	26.87
Paired Total Suspended Solids (mg/L)	34	23.41	15.78	31.04	14.50	0.00	88.00	6.00	40.00	478.01	21.86

Parameter	n	Mean	- 95% Confidence Interval	+ 95% Confidence Interval	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Variance	Standard Deviation
Probabilistic Total Solids (mg/L)	34	473.24	418.18	528.29	435.00	320.00	1200.00	390.00	500.00	24901.34	157.80
Paired Total Solids (mg/L)	34	448.53	394.13	502.92	410.00	320.00	1200.00	370.00	450.00	24303.83	155.90
Probabilistic Sulfate (mg/L)	34	54.54	46.60	62.47	50.50	8.20	100.00	37.00	72.00	516.91	22.74
Paired Sulfate (mg/L)	34	54.24	46.60	61.87	49.50	11.00	100.00	38.00	70.00	478.85	21.88
Probabilistic Temperature	34	22.80	21.41	24.19	22.87	12.49	30.66	20.46	24.90	15.84	3.98
Paired Temperature	34	22.74	21.24	24.23	22.94	12.02	31.27	20.02	24.92	18.32	4.28
Probabilistic TOC (mg/L)	34	4.11	3.21	5.01	3.55	1.30	14.00	2.20	5.30	6.64	2.58
Paired TOC (mg/L)	34	4.06	3.23	4.89	3.45	1.30	11.00	1.90	5.00	5.65	2.38
Probabilistic Turbidity (NTU)	34	55.54	-3.71	114.78	15.10	0.00	970.00	0.00	36.70	28833.88	169.81
Paired Turbidity (NTU)	34	28.83	17.03	40.64	19.80	0.00	153.10	0.00	45.00	1144.51	33.83
Probabilistic Zinc (ug/L)	34	20.26	14.86	25.67	18.50	0.00	71.00	12.00	27.00	240.14	15.50
Paired Zinc (ug/L)	34	20.65	14.95	26.35	17.50	0.00	57.00	11.00	33.00	266.96	16.34

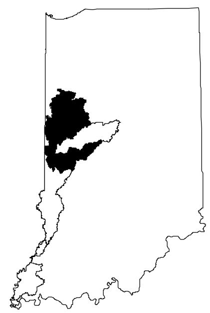
Middle Wabash River / Little Vermilion River Watershed HUC 05120108 1999 Paired Site Study

LEGEND

- County Boundaries
- # 1999 Probabilistic Sampling Sites
- # 1999 Paired Sampling Sites
- Rivers and Streams

Land Uses:

- Commercial and Services
- Confined Feeding
- Agricultural Lands
- Forest Lands
- Industrial
- Urban and or Built-Up
- Lakes and Reservoirs
- Residential
- Strip Mines
- Trans, Comm, Utl
- Transitional Areas
- Wetlands



Map Reference

Projection:	UTM Zone 16
Datum:	NAD 83
Printed:	February 2001
Plate Preparation:	Jim Butler

IDEM/Office of Water Quality
Assessment Branch/Surveys Section

1 : 550000

7 0 7 14 Miles

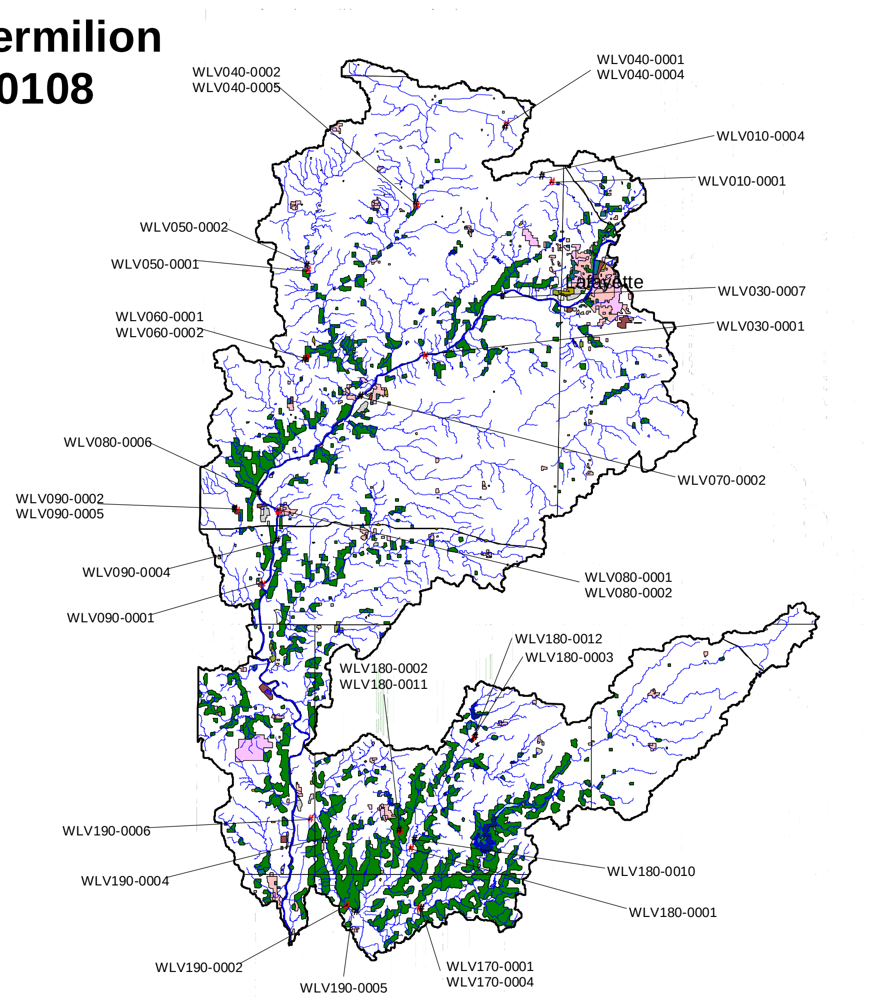
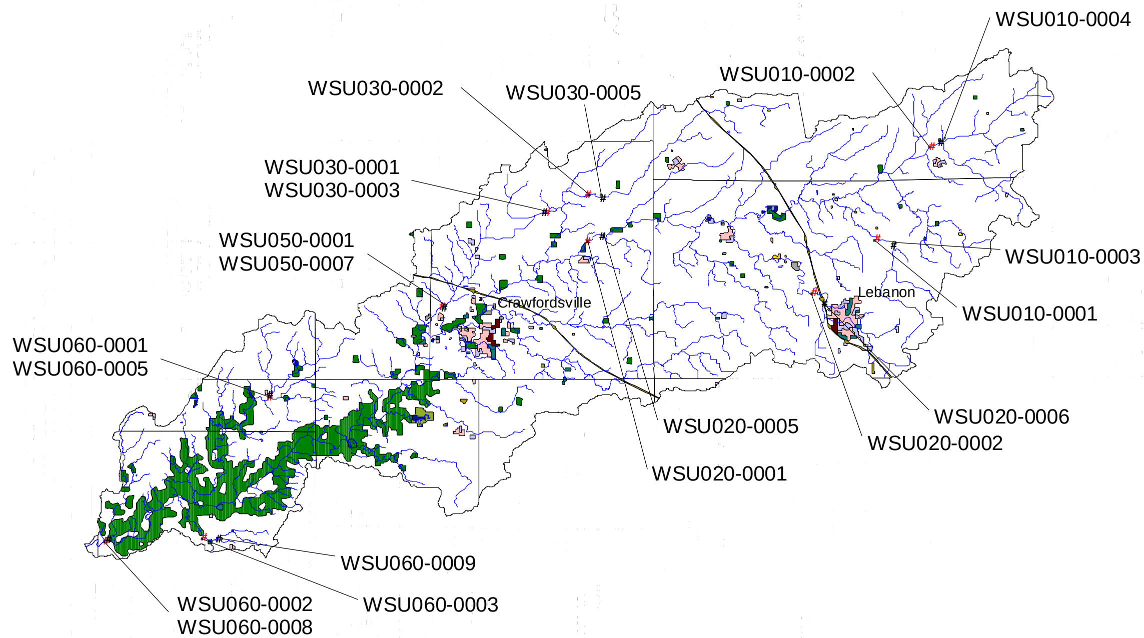


Plate 1. Middle Wabash River/Little Vermilion River Watershed Paired Study Sampling Locations

Indiana Department of Environmental Management (2001). *1999 Paired Site Study, Lower Wabash River Basin, Indiana* by Carl C. Christensen. Indiana Department of Environmental Management, Office of Water Quality, Assessment Branch, Surveys Section, Indianapolis, Indiana. IDEM 032/02/035/2001

Sugar Creek Watershed HUC 05120110
1999 Paired Site Study



LEGEND

- County Boundaries
- # 1999 Probabilistic Sampling Sites
- # 1999 Paired Sampling Sites
- Rivers and Streams

Land Uses:

- Commercial and Services
- Confined Feeding
- Agricultural Lands
- Forest Lands
- Industrial
- Urban and or Built-Up
- Lakes and Reservoirs
- Residential
- Strip Mines
- Trans, Comm, Utl
- Transitional Areas

1 : 400000

5 0 5 10 Miles

Map Reference

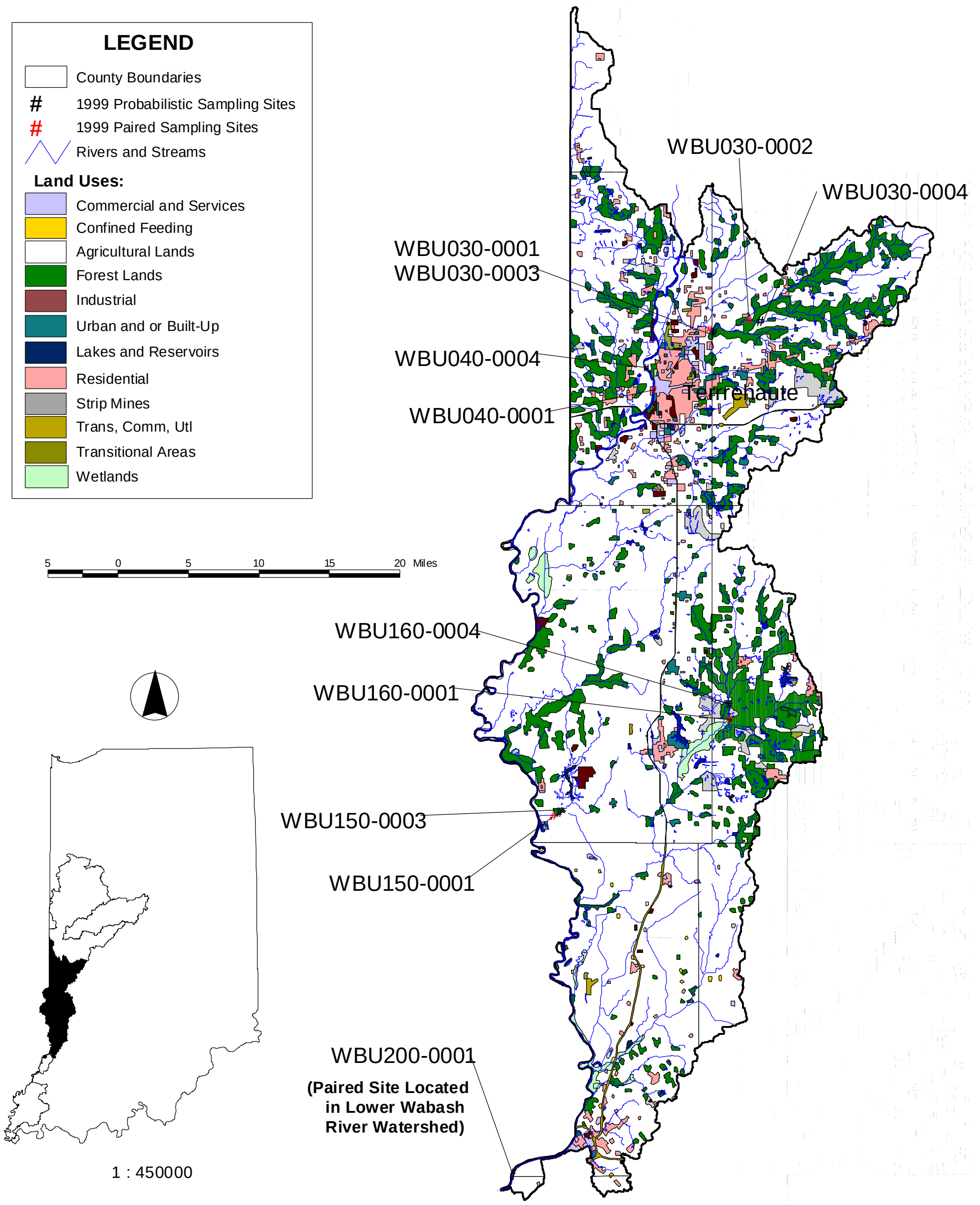
Projection:	UTM Zone 16
Datum:	NAD 83
Printed:	February 2001
Plate Preparation:	Jim Butler

IDEM/Office of Water Quality
Assessment Branch/Surveys Section

Plate 2. Sugar Creek Watershed Paired Study Sampling Locations

Indiana Department of Environmental Management (2001). *1999 Paired Site Study, Lower Wabash River Basin, Indiana* by Carl C. Christensen. Indiana Department of Environmental Management, Office of Water Quality, Assessment Branch, Surveys Section, Indianapolis, Indiana. IDEM 032/02/035/2001

Middle Wabash River/Busseron Creek Watershed
HUC 05120110 1999 Paired Site Study

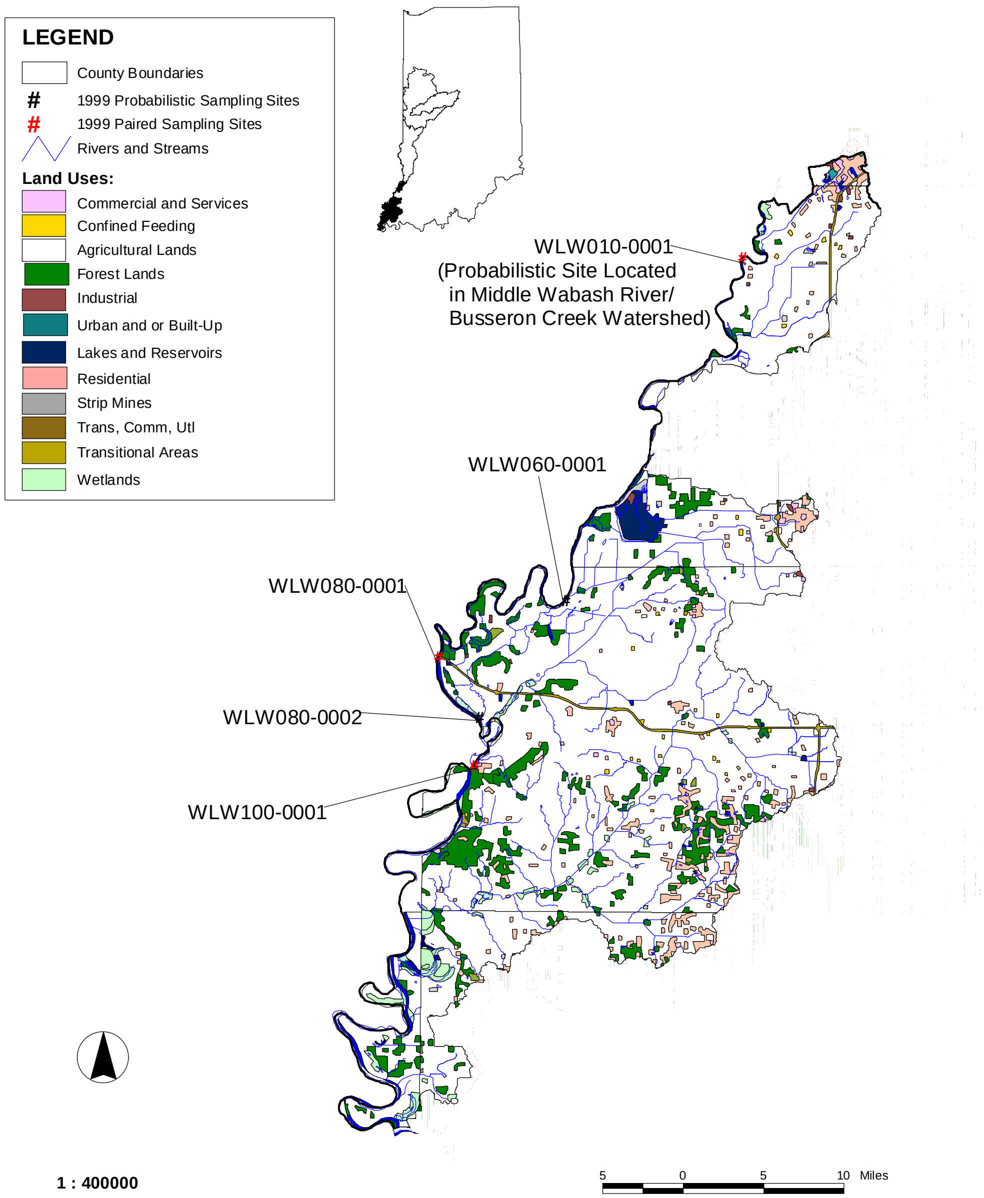


Map Reference	
Projection:	UTM Zone 16
Datum:	NAD 83
Printed:	February 2001
Plate Preparation:	Jim Butler

IDEM/Office of Water Quality
Assessment Branch/Surveys Section

Plate 3. Middle Wabash River/Busseron Creek Watershed Paired Study Sampling Locations
Indiana Department of Environmental Management (2001). *1999 Paired Site Study, Lower Wabash River Basin, Indiana* by Carl C. Christensen. Indiana Department of Environmental Management, Office of Water Quality, Assessment Branch, Surveys Section, Indianapolis, Indiana. IDEM 032/02/035/2001

Lower Wabash River Watershed
HUC 05120113 1999 Paired Site Study



Map Reference

Projection:	UTM Zone 16
Datum:	NAD 83
Printed:	February 2001
Plate Preparation:	Jim Butler

IDEM/Office of Water Quality
Assessment Branch/Surveys Section

Plate 4. Lower Wabash River Watershed Paired Study Sampling Locations

Indiana Department of Environmental Management (2001). *1999 Paired Site Study, Lower Wabash River Basin, Indiana* by Carl C. Christensen. Indiana Department of Environmental Management, Office of Water Quality, Assessment Branch, Surveys Section, Indianapolis, Indiana. IDEM 032/02/035/2001